**JOURNAL OF APPLIED INTELLIGENT SYSTEMS & INFORMATION SCIENCES**
VOL. 3, ISSUE 2, PP. 43-55, DECEMBER 2022.
*Available at:* www.journal.research.fanapsoft.com
**DOI:** https://doi.org/10.22034/JAISIS.2022.371239.1054

# ARTIFICIAL INTELLIGENCE AND CLINICAL DECISION MAKING: APPROACHES AND CHALLENGES

Nematollah Saeidi[1,*], Mahdi Torabi[2]

[1] *Department of Artificial Intelligence Engineering, University of Isfahan, Isfahan, Iran,*
[2] *Department of Computer and Electrical Engineering, University of Kashan, Kashan, Iran.*

## ABSTRACT

The use of artificial intelligence to target clinical problems has led to a revolution in clinical decision-making. Also, artificial intelligence has been feasible in most areas through the use of large labeled data and thanks to a significant increase in computing power and cloud storage. The success of these tools depends on understanding the intrinsic processes used along the conventional pathway through which clinicians make decisions. In the proposed work, we highlight the state-of-the-art in artificial intelligence and related approaches that can influence four levels: data acquisition, feature extraction, interpretation, and decision support. There are also clinical applications using artificial intelligence approaches widely which are explained here. In addition, many technical, medical, and ethical perspectives make the use of artificial intelligence open to criticism, thus some of its limitations and challenges related to regulations, explanation, validation, etc. are discussed here.

**KEYWORDS:** Artificial Intelligence, Decision-Making, Prediction, Clinical, Healthcare

## 1. INTRODUCTION

Computational support has gradually started to be used in clinical decision-making since the 1970s and 1980s. While this starting point was technically and clinically limited, rapid developments were progressively seen in the following years, facilitating the adoption of state-of-the-art approaches. In recent years, due to the application of medical and clinical data-based support systems, the implementation of data-based and knowledge-based systems in medical centers, hospitals, and healthcare centers has grown a lot, which is approved by the food and drug administration (FDA) (Adams et al., 1986; Adlung et al., 2021; Benjamens et al., 2020; Harish et al., 2021; Moja et al., 2014). Across the healthcare sector, artificial intelligence (AI) is being increasingly developed to streamline administrative processes, advance diagnostic, and therapeutic activities, and improve patient engagement. Alongside this growth, investment in healthcare AI companies continues, with a $2.5 billion investment in the first quarter of 2021. This trend is expected to continue, with up to three-quarters of all healthcare organizations expected to increase investment in AI-based solutions. One main application area is the integration of AI into clinical decision support systems to assist clinicians. Despite the flood of investment in health AI companies, widespread adoption of AI-enabled clinical decision support systems is limited. The experience of AI-powered clinical decision support systems highlights challenges facing wider adoption, such as insufficient evidence of clinical effectiveness, lack of standard methods for evaluating AI products, interoperability barriers, and limited capital to support (Rajkomar et al., 2019; Topol, 2019).

[*] *Corresponding Author, Email:* saeidi.n@eng.ui.ac.ir

**RECEIVED**: 22 NOVEMBER, 2022; **ACCEPTED**: 18 DECEMBER, 2022; **PUBLISHED ONLINE**: 26 DECEMBER, 2022

When decision support systems are gradually introduced into clinics, one of the most important principles is to evaluate the application and performance and carefully understand their limitations. Notably, the knowledge of these platforms should be extended to clinical users. AI can improve clinical decision-making in various ways by providing early warnings, facilitating diagnosis, performing extensive screening, personalizing treatment, and assessing patients' responses to treatment. An AI framework should be tested against clinical characteristics and current standard capacities, including clinician knowledge and experience. In some cases, AI models have proven to perform at least as well as clinicians. In recent years, proposed guidelines for use in AI-related clinical trials address issues by establishing frameworks for better uniformity in AI assessment (Montani & Striani, 2019; Sanchez-Martinez et al., 2022; Topol, 2019; Uddin et al., 2022).

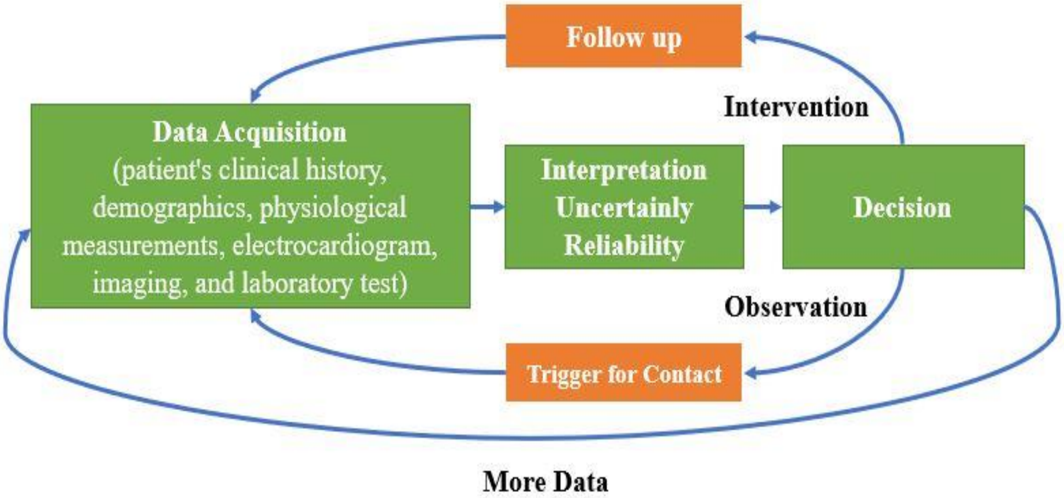


**Fig. 1.** General framework of clinical decision-making

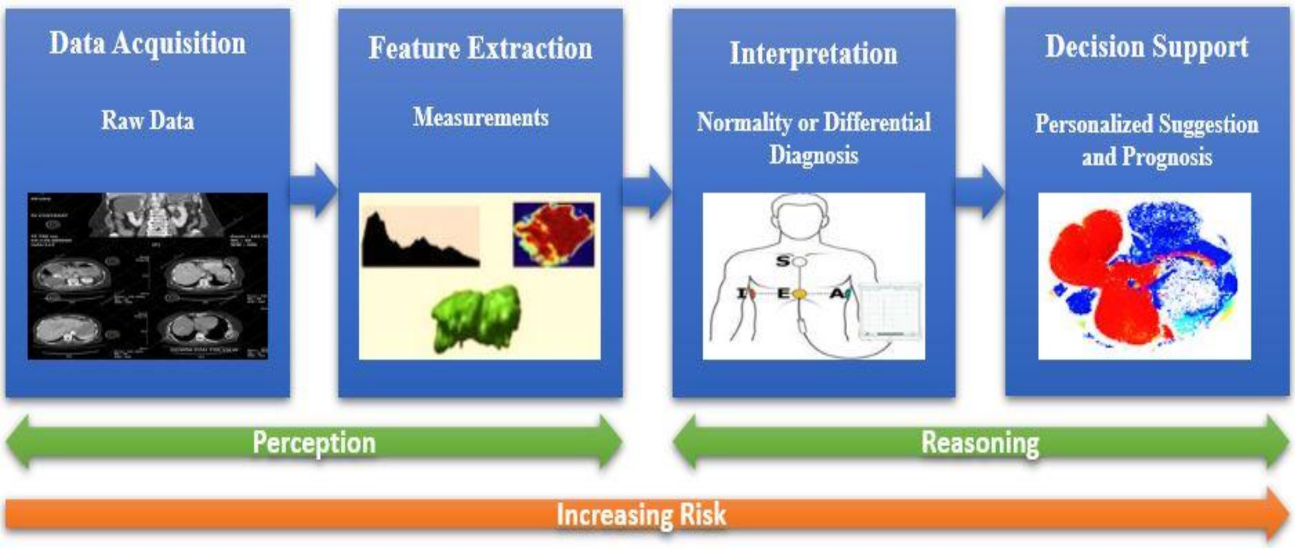


**Fig. 2.** Steps where AI supports clinical decision-making

Fig. 1 shows a general framework of clinical decision-making. It is formed by collecting data from patients, hospitals, and clinicians. In the next step, clinicians interpret the patient's condition by comparing it with the past information obtained from different people or instructions. This interpretation is based on data, AI models, and humans' inherent ability to contextualize information and recognize patterns. In addition, Clinicians assess the certainty of measurements and information and try to estimate how much the data can be relied upon. Afterward, doctors use the knowledge obtained from the patient's condition to decide. In addition to knowledge, developed AI models are essential factors in the decision-making process. This knowledge can include gathering more data and minimizing uncertainty in deciding and implementing an intervention (surgical treatment, drug, device) to improve the patient's condition (Harish et al., 2021).

In the evidence-based medicine era, millions of people are carefully examined, resulting in a considerable amount of complex and heterogeneous data. The use of algorithmic methods to analyze these data and strengthen clinical decision-making is made possible due to the ever-increasing computational power of AI. Big data leveraged by AI can provide clinicians with detailed information to present more accurate treatment diagnoses. They also can estimate the probabilities and costs of possible consequences. AI-enhanced decisions made by clinicians have the potential to improve outcomes, reduce care costs, and increase patient satisfaction.

Against Fig. 1, which presents the clinical decision-making process, Fig. 2 illustrates the tasks of this process according to the way AI can help. According to the studies, AI has shown a human-like performance in low-level tasks in which AI plays an essential role such as data acquisition and feature extraction. For high-level tasks such as patient status interpretation and decision support, AI enables the integration of complex and heterogeneous data into the decision-making process. However, these data are still immature and need to be validated. Taking clinical decision-making steps, from data acquisition, feature extraction, and interpretation to decision support increases the risk that directly affects patients and can cause irreparable risks for them in case of wrong decisions. Also, different models, approaches, or architectures do not disturb basic steps; these steps are the same for AI and can be known as a general process.

In addition to the four steps mentioned in Fig. 2, data normalization also can create an infrastructure that helps data sharing and aggregation. Data normalization is a hierarchy of data standardization that tends to increase data usability and reduce ambiguity and the obtained information of clinical is normalized to improve data sharing and care chain analysis. Medical professionals and clinicians identify data normalization as an effective and practical step for the cooperation of health information. They believe that normalization can ensure the software digest the data and achieves maximum efficiency when processing it. As the importance of efficient and optimal use of clinical personnel and its current epidemic has proven in medical communities, more appropriate normalization techniques help decision-makers prioritize patients with less error and allocate resources more efficiently and accurately. As fragmented or inconsistent data is not usually useful for AI models, any healthcare facility with the right resources can extract practical information from medical data. Therefore, choosing a suitable normalization technique is very important. Several normalization methods have been recently applied to transform heterogeneous input data into a dimensionless form. Decision makers can compare alternative options according to criteria with different scales mapping the decision matrix onto the interval [0,1]. The article (Vafaei et al., 2022) presents many normalization techniques and explains them. Among the introduced techniques, the most important ones include six primary techniques, which we have grouped into three linear, semi-linear, and non-linear categories. These techniques are known as ‘Max’, ‘Max-Min’, and ‘Sum’ for linear, ‘Vector’ for semi-linear, and ‘Logarithmic’ and ‘Fuzzification’ for non-linear.

In this paper, we present recent studies related to advanced AI technologies focusing on facilitating medical care. We highlight the characteristics of AI systems, describe an overall assessment of AI perspectives, and various types of learning, and critically discuss challenges in their application to clinical decision-making.

## 2. AI Perspectives

As AI methods and optimization of models can be used in different applications (Kurshan et al., 2020; Liu et al., 2021; Saeidi et al., 2019; Wang et al., 2020; Ying et al., 2018; Zhou et al., 2018), the speed of development of AI algorithms in the clinical and medical fields is gradually increasing. According to publications in recent years, there are significant deficiencies in health care causing its lack of up-to-date and rapid development. These deficiencies can include a large number of cases, including prediction errors, inefficiencies in workflow, treatment errors, loss of resources, and inequality. The AI era scientists believe that AI's traces are mitigating these deficiencies, and we can hear good news about this in the future. In the following, we examine the available evidence of AI in various clinical and medical aspects.

### *2.1. AI and data analysis*

One of the most important aspects of data analysis and AI in clinical problems is to understand how cancer evolves by applying transfer learning methods to multiregional tumor sequencing data and machine vision to analyze living cancer cells at single-cell resolution. These applications can ultimately help to improve decision-making conditions. Another example of AI is used to reconstruct neural circuits, which can provide a connectome understanding of electron microscopy. One of the most critical advances in AI in the field of medicine is understanding how the network of the human brain works. This achievement can lead to faster calculations and understanding of brain-machine interfaces in neuromorphic computing or reverse engineering of the brain to make computer chips. Another example is the development of human, animal, and machine behavior tracking through transfer learning methods and the use of medical and non-medical knowledge (Ameen et al., 2022a; Giordano et al., 2021; Lee et al., 2019).

Drug discovery based on AI can be mentioned in various applications such as conducting data mining in molecular structures, designing and manufacturing molecules, and predicting drug dosage for patients with different conditions. Also, as new activities, we can refer to the prediction of poisonings, data encryption in the pharmaceutical industry, and the discovery of drug interactions with the help of advanced AI algorithms. A well-known example is the Eve robot from the University of Cambridge and Manchester, which independently discovered an anti-malarial drug that is one of the ingredients of toothpaste. Such large companies and pharmaceutical startups can help the rapid development of this field (Williams et al., 2015).

### *2.2. AI and clinicians*

Soon, all clinicians, specialists, and medical organizations will resort to the practical use of AI and DL in their duties. These tasks mainly include analyzing and evaluating medical images (e.g., histopathology, x-ray, MRI, CT scan, ultrasound), ECG, endoscopy, and gastroscopy. As mentioned before, the use of AI algorithms has many applications in many clinical environments (Huss & Coupland, 2020; Kim, 2018). Applications in clinical settings can be mentioned in facilitating the diagnosis (stroke, autism, or electroencephalography) for neurologists, helping anesthesiologists to avoid low oxygen during surgery, diagnosing stroke or heart attack for paramedics, selecting live embryos for in vitro fertilization, and preventive surgery for breast cancer patients. Examples of the development of AI applications throughout the human lifespan are shown in Fig. 3. There is also a lot of effort among startups and technology companies to develop this field, including Google, Microsoft, Orbita, Robin Healthcare, Tenor.ai, and Sopris Health (Acosta et al., 2022; Ding et al., 2022; Wong et al., 2022; Zheng et al., 2021).

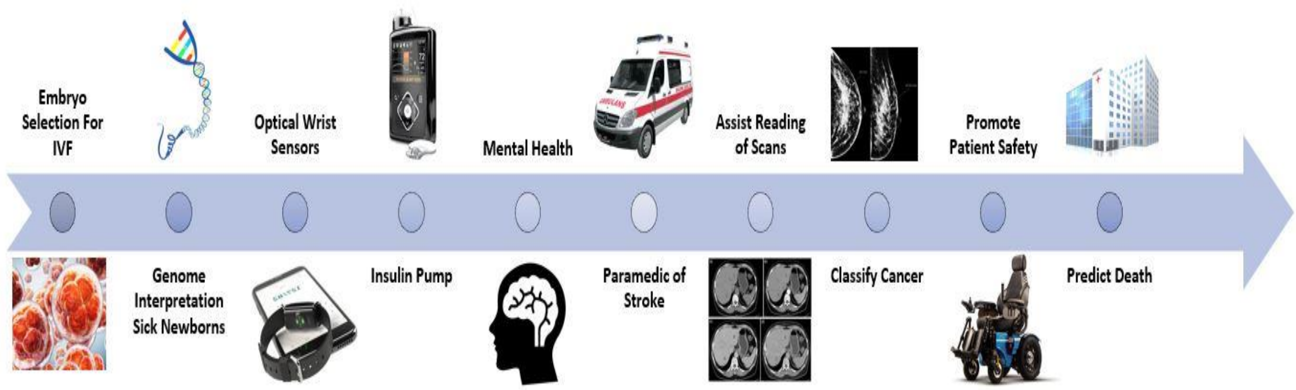


**Fig. 3.** The development of AI applications throughout the human lifespan

### *2.3. AI and health systems*

The possibility of predicting events such as lethal cancer in hospital palliative care can have a significant impact on its performance and improvement. For example, predicting the time people need to spend in hospital palliative care requires developing and expanding AI methods (Bhagwat et al., 2018; Elfiky et al., 2018; Makar et al., 2017; Miotto et al., 2016). Also, predicting the probability of a patient's hospitalization over time is an example showing the need for cutting-edge AI that can be achieved through data from electronic health records (EHR), AI, and DL approaches. Despite the need for further development in this field, many companies try to develop these methods. One such company is Careskore, which has developed systems predicting the risk of mortality based on available data. However, uncertainty among predictions is still an important aspect. When a model's AUC (Area under the ROC Curve) is 95%, it indicates a good model, but this cannot accurately predict a particular individual.

Although EHR data was discussed, imaging may also help to increase predictive accuracy. For example, many studies have been conducted to determine and predict biological age, and the result has shown that this task can be conducted using biomarkers based on DNA methylation. In this case, there is a point in the accuracy of the prediction. Because a large part of input data is unstructured and incomplete (or data drawn from example a socioeconomic, behavioral, biological, and physiological sensor) due to clinician notes in medical files, it is impossible to enter these inputs into AI algorithms. In addition, some data are known as challenges due to having few samples (Daunay et al., 2019). In the end, the AUC is known as a metric in this field, but this metric cannot use sensitivities and characteristics that are of interest to clinicians. Therefore, the impact of AI in the healthcare environment is not yet clear and is impossible until there will be a criterion with strong validation based on statistical analysis in real clinical settings.

According to the three AI perspectives raised, there is a possibility of high development due to big data in AI and DL. This unique opportunity can manifest in three medical perspectives: data, clinicians, and health systems. The existence of real data will be one of the essential items for the development of any clinical application. For clinicians, it can facilitate standard and efficient data analysis to speed up patient prediction and improvement and create personalized medical guidelines. For health systems, it improves circulation, speeds up work, and minimizes medical errors. The existence of big data requires expertise in several medical fields to be able to combine them into a large productive collection. The ever-increasing development of high-throughput technologies and the implementation of EHRs led to the rapid growth of electronic health record data. Analyzing multimodal data that includes complete medical and clinical information can extract helpful knowledge according to the relevant clinical tasks. Mainly, EHR data faces the challenge of high dimensions in the data, which also results in the need for high computing power. One of the ways to face the upcoming challenge of EHR data is dimension reduction methods without losing valuable information. Another challenge is privacy, personal autonomy, trust, and fairness, the ethical and legal challenges are sensitive issues that need to be focused on by big data in healthcare. Also, among other challenges are technical and infrastructure issues that can endanger health care. Factors of this challenge are mentioned due to analytical flows in data review, data protection, data heterogeneity, and lack of proper infrastructure for data storage.

The explanation presented in more detail is shown in Fig. 4. In Fig. 4, the general steps including data source, data, data storage, analytics, and improved outputs, are separately illustrated. Collected data from medical research and hospitals, including clinical, public health records, sensing, EHR, and OMICS were stored by master person index (MPI) and operational data store (ODS). The clinical data warehouse also enables healthcare organizations to evaluate disease management programs that directly and indirectly affect the quality of patient well-being. Furthermore, the values of diagnostics, prescriptive, descriptive, and predictive, are analyzed to improve outputs. Among the outputs that can be improved, many items are related to risk and disease management, items leading to reducing fraud, medical imaging, personalized care improvement, new therapies development, population health, cost reduction, medication abuse reduction, medical research, precision medicine, preventive medicine, and patient's engagement enhancement (Sahu et al., 2022).

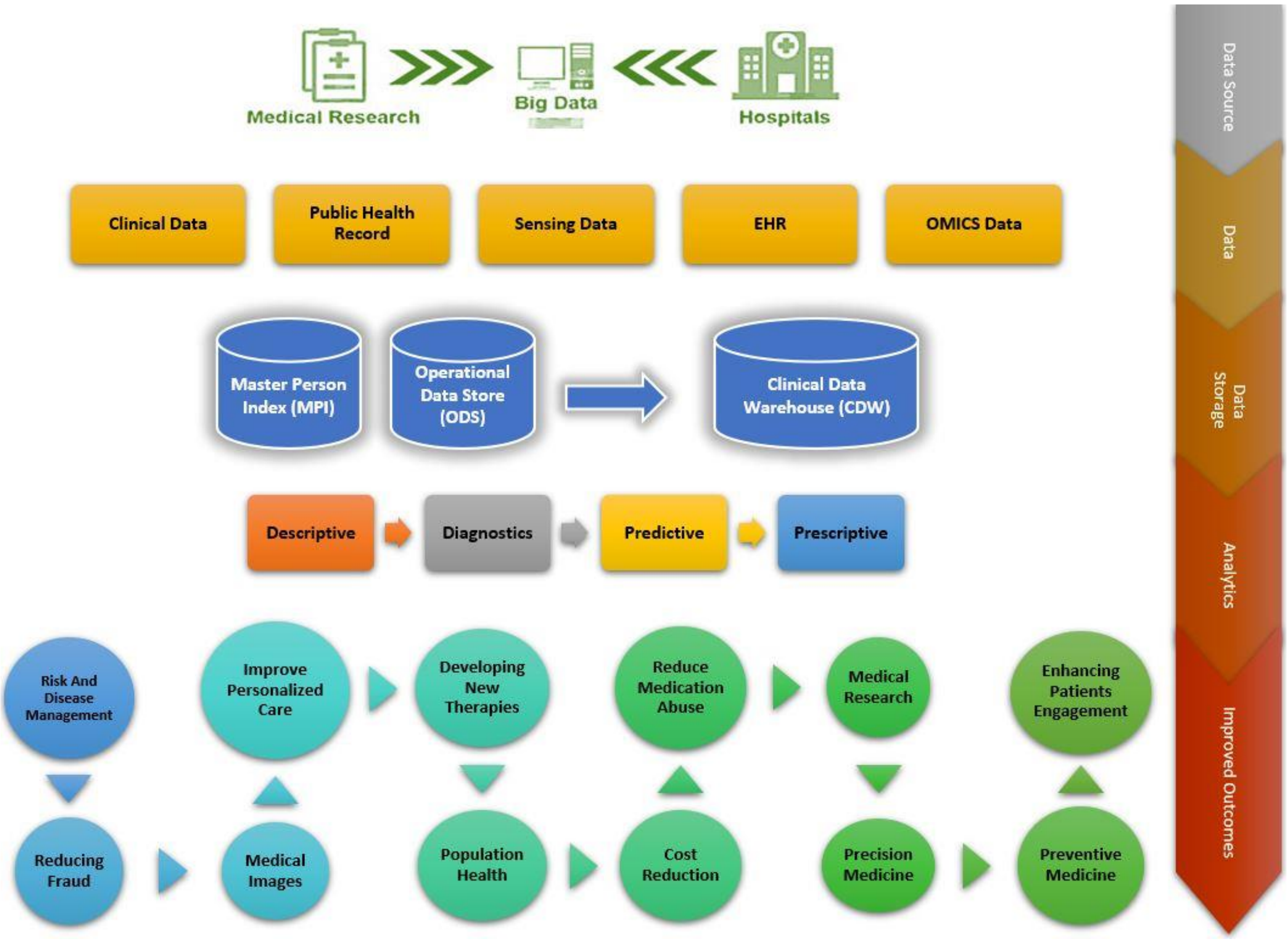


**Fig. 4.** Improvement of Data in Healthcare and Improvement Outcomes

## 3. Various Types of Learning

In general, to enable clinical applications, AI approaches in this field can be classified into the following methods:

### *3.1. Unsupervised learning*

In this model, there is no label for input information. The methods aim to discover and reason about complex structures in the data. Methods such as FA, PCA, and autoencoder try to find low-dimensional data understanding of the relationships between data features. Reciprocally, mixture modeling and cluster analysis identify groups in the data. As an example, the aim of Le et al. (2022) is to achieve a method to deal with sparsity by compressing the feature space of clinical presentation, where limited clinical notes can be effectively dealt with. The mentioned method applies to an autoencoder model to use sparsity reduction in clinical note representation. The first motivation of this work was to determine how to compress sparse and high-dimensional data by reducing the dimensions of the clinical note representation feature space. In another method (Chushig-Muzo et al., 2021), an AE-based method has been presented by combining probabilistic models based on Gaussian mixture models and hierarchical clustering supported by Kullback Leibler divergence. Also, for clinical validation, it used real data from the EHR of the University Hospital of Fuenlabrada in Spain. The obtained results showed that this method could perform clustering well and find groups of patients with similar health conditions with the help of patterns related to diagnosis and drug codes (Adlung et al., 2021).

### 3.2. *Supervised learning*

Unlike unsupervised models, labeled data is required to learn the mapping between input and output data. This method is widely used in clinical decision-making systems, where each model maps labeled input data onto output using a function. After training, the mentioned function can predict the output with the new input data. Supervised learning is primarily used in regression, where the output variable is continuous, and in classification, where the output variable is discrete. Examples of the most well-known supervised methods are artificial neural networks (ANN), decision trees, linear regression (LR), and support vector machines (SVM). Also, SVM and LR are easy-to-interpret models that look for linear relationships between input data and outcome.

A decision tree also recursively divides the entire feature space into smaller domains during training and associates each smaller domain with an outcome value according to the training data it holds. Given an unlabeled sample, this tree determines smaller domains corresponding to the test data according to the learned partitioning and outputs the value corresponding to this smaller domain. When there are no constraints, decision trees may partition the feature space in such a way that each smaller domain contains only one data and will perform poorly if there are examples outside the training data. In contrast, shallow decision trees, which allow only a small number of domains, miss complex structures and subtle relationships in the data. Also, ensemble methods such as random forests, AdaBoost, and XGBoost solve such problems by combining several weak models to create a diverse and accurate model. In addition to previous models, transforming data into a latent feature space by ANNs, a simple regression is used for the latent representation and returns accurate outcomes. Learning this latent representation is performed by applying a sequence of linear mappings, where a non-linear activation function follows each linear transformation to allow for composite transformations. Each pair of linear and activation functions results in an intermediate representation of the data related as a layer. Also, Fitting parameters traditionally associated with ANNs need a training dataset that is large enough to avoid data memorization (Patrício et al., 2022; Uddin et al., 2022).

### 3.3. *Weakly supervised learning*

This model attempts to build predictive models by learning with weak supervision. Due to the particular requirement of labeled datasets in the supervised model to have good results, such models deal with this challenge by using limited, imprecise, and noisy data annotations for supervised approaches. For instance, A task of weakly supervised learning was conducted to learn and identify cancerous lesions from a dataset for labeling histopathology images applying multiple-instance learning.

For example, this approach (Hu et al., 2020) proposed a weakly supervised learning method for classifying COVID-19 infection in CT images. This method tries to improve the manual labeling of CT images but can still detect infection and accurately distinguish COVID-19 from other images. Based on the results obtained, we can have the growth and development of the application of the implemented technique on a wide and large scale in medical and clinical studies. In another paper (Ouyang et al., 2020), the abnormality localization was considered for medical and clinical tasks. Despite the stunning development and growth of current methods and having high diagnostic accuracy, clinicians do not fully trust these algorithms for diagnostic decision-making purposes due to the lack of comprehensive decision-making reasoning. In the same way, special attention was paid to using a new attention-driven weakly supervised algorithm, which includes a hierarchical attention mining framework that integrates activation-based and gradient-based visual attention. The method outlined significant localization performance improvement evaluating a large-scale chest X-ray dataset.

### 3.4. *Self-supervised learning*

This model is the process in which it trains its model to learn one part of the input from another part of the input. In (Xue et al., 2020), a model was trained for the self-supervised task of an unlabeled EHR dataset and then tuned for the supervised task of estimating time-to-event for scenarios such as death and kidney failure using a smaller labeled dataset. This pre-training technique is a remarkable example of the transfer learning method, which was known as a solution to the lack of large-scale data. Transfer learning uses the knowledge of pre-trained models to train a new model or extend the capabilities of existing models. Also, self-supervised learning can be helpful in histopathology image segmentation for cancer diagnosis and assess severity levels of diabetic retinopathy from funduscopic images. A recent review paper by Krishnan et al. (2022) described self-monitoring learning approaches used in healthcare to develop models that use multivariate data sets and the challenges of unbiased collection. With self-supervised learning that considers unlabeled data, this method can obtain the required information about a specific data source (e.g., medical images and signals) using unlabeled data. This approach is recognized as a cutting-edge starting point even beyond medicine and healthcare due to the lack of the need to collect large data sets. In another work, the authors were able to recognize diseases on chest radiographs and classify them at the level of radiologists into different types of pathology, even with unannotated examples of diseases on chest images. Also, Li et al. (2020) reported the importance of analyzing and investigating different retinal diseases from fundus images for clinical decision-making. Developing these approaches is challenging due to the need for large amounts of human data. Because diagnosing different types of retinal diseases greatly benefits from other imaging approaches, this article implemented a new retinal disease diagnosis method for learning self-monitoring features using multimodal data.

### *3.5. Reinforcement learning*

This model is based on reward and punishment in that it is trained to perform tasks by receiving rewards for good actions and punishment for bad ones. An example of reinforcement learning is longitudinal evaluations of treatment options for patients, where the model recommends the drug with the reward of a good clinical outcome. Deep reinforcement learning methods are also based on popular topics such as the Q-learning and Markov decision process adapted to neural networks.

Liu et al. (2019) presented a review paper on deep reinforcement learning and clinical decision-making. Their work describes deep reinforcement learning with neural networks and investigates the most important approaches, developments, and challenges. Finally, it reviews the most important applications of clinical reinforcement learning for clinical decision-making systems. Such applications have demonstrated that deep reinforcement learning can be applied to recommend treatment for various applications with different data sources, including genetic data, disease databases, and EHR. Applications contain constructing clinical motifs from clinical notes, settings and duration of mechanical ventilation, dosage for patients with acute or chronic disease, and medication/fluid choice. As practical examples, deep reinforcement learning was used by Nemati et al. (2016) to optimize drug dosage. Also, Prasad et al. (2017) used a reinforcement learning method to wean mechanical ventilation in the intensive care unit. Despite developing advanced methods and approaches to optimizing patient care, safety and accountability are among the highest priorities. The leaders' approaches may be dynamic and personal, but the logic, decision-making, and performance behind them can be questionable. Another point is that it is more challenging to use reinforcement learning in clinical environments compared to simulation environments (Giordano et al., 2021; Haarnoja et al., 2017; Nemati et al., 2016).

## 4. CHALLENGES AND LIMITATIONS

This section discusses the most significant challenges that may appear when dealing with any clinical problem with AI approaches.

### *4.1. Explainability*

AI methods for regression and classification are related to data that are robust and reliable enough to make predictions from previously unseen data. AI model explanation based on the underlying data only refers to the statistical sense, where measuring the amount of variance in the output labels explained by the pre-trained model is possible. Accordingly, the ability to explain AI methods should not be confused with causal inference or observations. A decision tree as a simple example of a tree-like model is formed by answering different questions. Also, linear regression, where a coefficient's absolute value and sign are assigned to a predictor, describes its strong effect (Kovalchuk et al., 2022).

### *4.2. Causality*

This challenge discusses criteria for the quality of explanations presented by AI methods. In clinical decision-making, medical professionals need specific criteria to understand why an AI algorithm produces a specific result. This process was facilitated by interactive AI, where a human domain expert gives knowledge for the learning process. A thorough understanding of causality is the foundation for a rational attribution of AI-based medical discoveries in compliance with regulatory barriers.

### *4.3. Data*

Due to increasing public access to clinical data, high-quality data are still limited. Especially when we have rare medical cases, clinical data are rarely available. A key example is the EHR data collection challenge, where integrating data from multiple sources is challenging despite the data's different formats and complex structures. When using inadequate data to train an AI model, there is a risk of selecting random correlations that appear due to sampling size or bias, which have nothing to do with true relationships between involved parameters. Also, models are prone to poor generalizability and provide inaccurate predictions. A related challenge is the public sharing of data and code that enables practical implementation of the models, which is frustratingly limited. According to the mentioned considerations, the emphasis is on sharing to enable validation, collaboration, and comparison of their results between researchers and study populations. In addition, most of the time, medical data are stored separately in the systems, which makes the level of access difficult and makes the comparison at the population level practically tough and almost impossible. EHRs contain unstructured data and information that healthcare providers and clinical researchers cannot use. In this regard, machine learning models can organize this data or be used to directly integrate complex unstructured data for high-throughput phenotyping to identify patient groups (Ameen et al., 2022a; Kelly et al., 2019).

### *4.4. Security*

A fundamental issue of the future of medicine AI is how to ensure data privacy and security. Given the pervasive problems of hacking and data breaches, there is little interest in using algorithms that risk exposing details of a patient's medical history (Topol, 2019). Since DL models require large datasets for training, machine learning raises several data security and privacy issues. In this regard, the most secure way to transfer data between healthcare organizations is still vague, and stakeholders no longer underestimate the risks of losing sensitive data. Hacking is one of the most critical issues because hackers can manipulate a decision-making model to cause damage on a large scale. A practical solution is Blockchain, which enables data exchange systems that are cryptographically secure and irreversible by providing a public and immutable log to regulate data access. However, blockchain technology has negative points, such as being slow, expensive to maintain, and difficult to scale. In addition, federated learning (an algorithm on decentralized edge servers that keeps local data samples without exchanging them) enables updating a learning model without sharing individual information with a central system and can ensure patient data security. New models of health data ownership with individual rights, the use of highly secure data platforms, and government regulation are needed to address evolving security issues. Otherwise, the chance of progress in AI for medicine will stop (Sanchez-Martinez et al., 2022).

### *4.5. Generalizability*

Typically, an AI system performs better on newly collected data than on similar sources. Generalizability measures how useful a model will be in real-life data. For example, suppose an AI system is trained and tested on data from one hospital and presents poorly on data from another hospital. In that case, the AI system may have less generalizability. AI systems with higher generalizability are often preferred for clinical usage. There are several reasons for the expected generalizability of a trained model. However, in the absence of a precise definition of this concept, new guidelines are generally provided that require validation (Sanchez-Martinez et al., 2022).

### 4.6. Trials in real-life data conditions

Determining whether an AI system is ideal for clinical utilization is currently one of the most challenging issues in clinical decision-making systems. Mostly AI systems are validated in retrospective or prospective experiments and some cases with simulated data. As these trials can be helpful to proof-of-concept methods, more trials are needed to allow better evaluation of the AI system. Although trials may provide strong evidence for the performance of an AI system, they may face interpretive difficulties and challenges due to the lack of fine-grained controls (Adlung et al., 2021; Topol, 2019). Recently, while closely monitoring the deployment of an AI system, a team from google research observed many unexpected issues they never encountered in the clinic, which sometimes prevent diagnosis and add unexpected costs. Therefore, This emphasizes further research and investigation for the AI system to face life obstacles in real conditions, which mainly requires adjustments in the model or how it is deployed in hospitals, clinics, and research centers.

### 4.7. Regulatory

Using AI for clinical decision-making inevitably raises legal challenges regarding medical negligence arising from learning failures. When such negligence occurs, the legal system should guide what entity is responsible for which recommendations were made. Additionally, the evolving nature of AI methods presents a unique challenge for regulatory agencies, and how best to evaluate updates remains unclear. Regulatory constraints are often encountered in the certification process or AI-based software for use in clinical trials or during adaptation to a real-world environment. Although legal considerations vary significantly between global regions, the concept of learning in change poses a significant regulatory challenge in most regulatory settings. Ethical considerations also constitute an important unresolved issue using predictive AI pipelines. For instance, to share data, these types of requests must comply with data protection considerations. Implications of personal health predictions may have significant medical and economic consequences (Sanchez-Martinez et al., 2022).

### 4.8. Causal AI vs. Predictive AI

Predictive AI based on the correlation of input data and results may not be sufficient to impact the healthcare system. Indeed, this form of learning can be misleading if important causal variables are not analyzed. According to the above-mentioned points, the main objective is to find the main reason why the algorithm decides to reach that decision. Questions are addressed by causal AI, a robust analysis aimed at inferring the mechanisms of the system generating the output data. Therefore, causal models can provide detailed maps of variable interactions with the result that users can simulate future cause and effect (Topol, 2019).

### 4.9. Validation

In general, if an algorithm outperforms humans in predictive tasks, extensive validation should be mandatory. One of the essentials of AI algorithms for deployment in hospitals is to improve financial and patient outcomes. Validation should be through multicenter randomized prospective trials to assess whether the trained models can be applied elsewhere or not. In this regard, examples of prospective AI trials that were evaluated in a real clinic are rare. Among them, one of the most significant advantages of AI models is the improvement of performance by having more data. However, this is a challenge, particularly for neural networks with catastrophic forgetting (the tendency to completely and suddenly forget learned information in new learning). In addition, retraining the entire dataset is time-consuming and resource-consuming. One way could be federated learning, which runs models locally to improve them with a user's data.

## 5. Conclusions and Future Directions

The rapid growth of publications and extensive participation of laboratories in the clinical and medical fields has significantly impacted clinical decision-making systems. However, maintaining and continuing the process is also a condition for progress. AI systems are gradually advancing from established clinical scenarios to multiple domains. The widespread and unique involvement of AI systems in clinical decision-making may have a significant impact but it requires continued extensive research. Another essential point is clinical evaluation, carried out by clinical performance metrics to ensure the performance of implemented approaches based on AI. Among the evaluation factors, it is possible to understand the effects of AI on the performance of care by examining the diversity of healthcare professionals and how it works. Clinical professionals should keep in mind that the design and implementation of AI services should be highly effective for a global community. This is necessary to improve the interpretability of human-algorithm interactions for their acceptance by developing thoughtful monitoring approaches.

Research to define hybrid systems that can integrate knowledge-based and data-driven methods is currently underway. Such approaches require the combination of data and knowledge to enable maximum use of knowledge. Also, researchers believe this path will be one of the pioneers of a critical leap in the future of decision-based systems. This shows that data-based approaches can benefit from knowledge-based approaches and their generalization and abstraction capabilities. Also, data and knowledge provide valuable explainable decision support. Despite all the empirical research that has been done, the essential need is to demonstrate improvement in patient outcomes or processes of care. The emerging regulatory guidelines can help clinicians, startups, companies, and high-tech healthcare organizations choose the most appropriate paths in developing research that can lead to clinical improvement.

Recently, widespread research has been led to develop (1) cutting-edge human-based methods, (2) cloud computing infrastructure to support the storage and retrieval of extensive volumes of structured and unstructured data, and (3) the acceleration of computational processing power in both domains cloud computing and quantum computing has been done (Ameen et al., 2022b). For most AI researchers, it is believed that the union of these three fields will lead to the maximum growth of AI research, called artificial general intelligence. Therefore, it's clear that AI systems are here to stay, and everyone will likely have to grapple with the ethical issues of patient autonomy in the presence of currently inexplicable AI systems. It's worth noting that from the epistemological point of view, many differences can be found in how AI and humans deal with the facts they want to interpret in the data. A big gap in our understanding of AI and the understanding of AI in the world can cause this. Now this question arises if AI is consistently correct in longitudinal observations, is there a threshold limit to trust AI in favor of human judgment? This question is an open question that has not yet been answered. According to EU regulators, a system that can't explain itself, we don't understand that system doesn't apply in clinical practice. This view is still correct. As more evidence becomes available shortly, continuing this position may be an ethical danger.

## References


Acosta, J. N., Falcone, G. J., Rajpurkar, P., & Topol, E. J. (2022). Multimodal biomedical AI. *Nature Medicine*, *28*(9), 1773–1784. https://doi.org/10.1038/s41591-022-01981-2

Adams, I. D., Chan, M., Clifford, P. C., Cooke, W. M., Dallos, V., de Dombal, F. T., Edwards, M. H., Hancock, D. M., Hewett, D. J., & McIntyre, N. (1986). Computer-aided diagnosis of acute abdominal pain: a multicentre study. *BMJ*, *293*(6550), 800–804. https://doi.org/10.1136/bmj.293.6550.800

Adlung, L., Cohen, Y., Mor, U., & Elinav, E. (2021). Machine learning in clinical decision making. *Med*, *2*(6), 642–665. https://doi.org/10.1016/J.MEDJ.2021.04.006

Ameen, S., Wong, M.-C., Yee, K.-C., & Turner, P. (2022a). AI and Clinical Decision Making: The Limitations and Risks of Computational Reductionism in Bowel Cancer Screening. *Applied Sciences*, *12*(7). https://doi.org/10.3390/app12073341

Ameen, S., Wong, M. C., Yee, K. C., & Turner, P. (2022b). AI and Clinical Decision Making: The Limitations and Risks of Computational Reductionism in Bowel Cancer Screening. *Applied Sciences 2022, Vol. 12, Page 3341*, *12*(7), 3341. https://doi.org/10.3390/APP12073341

Benjamens, S., Dhunnoo, P., & Meskó, B. (2020). The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *NPJ Digital Medicine*, *3*, 118. https://doi.org/10.1038/s41746-020-00324-0

Bhagwat, N., Viviano, J. D., Voineskos, A. N., Chakravarty, M. M., & Initiative, A. D. N. (2018). Modeling and prediction of clinical symptom trajectories in Alzheimer's disease using longitudinal data. *PLOS Computational Biology*, *14*(9), 1–25. https://doi.org/10.1371/journal.pcbi.1006376

Chushig-Muzo, D., Soguero-Ruiz, C., de Miguel-Bohoyo, P., & Mora-Jiménez, I. (2021). Interpreting clinical latent representations using autoencoders and probabilistic models. *Artificial Intelligence in Medicine, 122,* 102211. https://doi.org/https://doi.org/10.1016/j.artmed.2021.102211

Daunay, A., Baudrin, L. G., Deleuze, J.-F., & How-Kit, A. (2019). Evaluation of six blood-based age prediction models using DNA methylation analysis by pyrosequencing. *Scientific Reports*, *9*(1), 8862. https://doi.org/10.1038/s41598-019-45197-w

Ding, K., Zhou, M., Wang, Z., Liu, Q., Arnold, C. W., Zhang, S., & Metaxas, D. N. (2022). Graph Convolutional Networks for Multi-modality Medical Imaging: Methods, Architectures, and Clinical Applications. *ArXiv*, *abs/2202.0*.

Elfiky, A. A., Pany, M. J., Parikh, R. B., & Obermeyer, Z. (2018). Development and Application of a Machine Learning Approach to Assess Short-term Mortality Risk Among Patients With Cancer Starting Chemotherapy. *JAMA Network Open*, *1*(3), e180926–e180926. https://doi.org/10.1001/jamanetworkopen.2018.0926

Giordano, C., Brennan, M., Mohamed, B., Rashidi, P., Modave, F., & Tighe, P. (2021). Accessing Artificial Intelligence for Clinical Decision-Making. *Frontiers in Digital Health*, *3*. https://doi.org/10.3389/fdgth.2021.645232

Haarnoja, T., Tang, H., Abbeel, P., & Levine, S. (2017). Reinforcement Learning with Deep Energy-Based Policies. *Proceedings of the 34th International Conference on Machine Learning - Volume 70*, 1352–1361.

Harish, V., Morgado, F., Stern, A. D., & Das, S. (2021). Artificial Intelligence and Clinical Decision Making: The New Nature of Medical Uncertainty. *Academic Medicine : Journal of the Association of American Medical Colleges*, *96*(1), 31–36. https://doi.org/10.1097/ACM.0000000000003707

Hu, S., Gao, Y., Niu, Z., Jiang, Y., Li, L., Xiao, X., Wang, M., Fang, E. F., Menpes-Smith, W., & Xia, J. (2020). Weakly supervised deep learning for covid-19 infection detection and classification from ct images. *IEEE Access*, *8*, 118869–118883.

Huss, R., & Coupland, S. E. (2020). Software-assisted decision support in digital histopathology. *The Journal of Pathology*, *250*(5), 685–692. https://doi.org/10.1002/path.5388

Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, *17*(1), 1–9.

Kim, J. T. (2018). Application of Machine and Deep Learning Algorithms in Intelligent Clinical Decision Support Systems in Healthcare. *Journal of Health and Medical Informatics*, *09*.

Kovalchuk, S. V, Kopanitsa, G. D., Derevitskii, I. V, Matveev, G. A., & Savitskaya, D. A. (2022). Three-stage intelligent support of clinical decision-making for higher trust, validity, and explainability. *Journal of Biomedical Informatics*, *127*, 104013. https://doi.org/10.1016/j.jbi.2022.104013

Krishnan, R., Rajpurkar, P., & Topol, E. J. (2022). Self-supervised learning in medicine and healthcare. *Nature Biomedical Engineering*. https://doi.org/10.1038/s41551-022-00914-1

Kurshan, E., Shen, H., & Yu, H. (2020). Financial Crime Fraud Detection Using Graph Computing: Application Considerations Outlook. *Proceedings - 2020 2nd International Conference on Transdisciplinary AI, TransAI 2020*, 125–130. https://doi.org/10.1109/TRANSAI49837.2020.00029

Le, T.-D., Noumeir, R., Rambaud, J., Sans, G., & Jouvet, P. (2022). Adaptation of Autoencoder for Sparsity Reduction From Clinical Notes Representation Learning. *ArXiv Preprint ArXiv:2209.12831*.

Lee, K., Turner, N., Macrina, T., Wu, J., Lu, R., & Seung, H. S. (2019). Convolutional nets for reconstructing neural circuits from brain images acquired by serial section electron microscopy. *Current Opinion in Neurobiology*, *55*, 188–198. https://doi.org/10.1016/j.conb.2019.04.001

Li, X., Jia, M., Islam, M. T., Yu, L., & Xing, L. (2020). Self-Supervised Feature Learning via Exploiting Multi-Modal Data for Retinal Disease Diagnosis. *IEEE Transactions on Medical Imaging*, *39*(12), 4023–4033. https://doi.org/10.1109/TMI.2020.3008871

Liu, S., Ngiam, K. Y., & Feng, M. (2019). Deep Reinforcement Learning for Clinical Decision Support: A Brief Survey. *ArXiv*, *abs/1907.0*.

Liu, S., Ni'mah, I., Menkovski, V., Mocanu, D. C., & Pechenizkiy, M. (2021). Efficient and effective training of sparse recurrent neural networks. *Neural Computing and Applications 2021 33:15*, *33*(15), 9625–9636. https://doi.org/10.1007/S00521-021-05727-Y

Makar, M., Oh, J., Fusco, C., Marchesani, J., McCaffrey, R., Rao, K., Ryan, E. E., Washer, L., West, L. R., Young, V. B., Guttag, J., Hooper, D. C., Shenoy, E. S., & Wiens, J. (2017). A data-driven approach to predict the daily risk of Clostridium difficile infection at two large academic health centers. *Open Forum Infectious Diseases*, *4*(suppl_1), S403–S404. https://doi.org/10.1093/ofid/ofx163.1009

Miotto, R., Li, L., Kidd, B. A., & Dudley, J. T. (2016). Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. *Scientific Reports*, *6*(1), 26094. https://doi.org/10.1038/srep26094

Moja, L., Kwag, K. H., Lytras, T., Bertizzolo, L., Brandt, L., Pecoraro, V., Rigon, G., Vaona, A., Ruggiero, F., Mangia, M., Iorio, A., Kunnamo, I., & Bonovas, S. (2014). Effectiveness of Computerized Decision Support Systems Linked to Electronic Health Records: A Systematic Review and Meta-Analysis. *American Journal of Public Health*, *104*(12), e12–e22. https://doi.org/10.2105/AJPH.2014.302164

Montani, S., & Striani, M. (2019). Artificial Intelligence in Clinical Decision Support: a Focused Literature Survey. *Yearbook of Medical Informatics*, *28*(1), 120–127. https://doi.org/10.1055/S-0039-1677911

Nemati, S., Ghassemi, M. M., & Clifford, G. D. (2016). Optimal medication dosing from suboptimal clinical examples: a deep reinforcement learning approach. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, *2016*, 2978–2981. https://doi.org/10.1109/EMBC.2016.7591355

Ouyang, X., Karanam, S., Wu, Z., Chen, T., Huo, J., Zhou, X. S., Wang, Q., & Cheng, J.-Z. (2020). Learning hierarchical attention for weakly-supervised chest X-ray abnormality localization and diagnosis. *IEEE Transactions on Medical Imaging*, *40*(10), 2698–2710.

Patrício, C., Neves, J. C., & Teixeira, L. F. (2022). *Explainable Deep Learning Methods in Medical Imaging Diagnosis: A Survey*. https://doi.org/10.48550/arxiv.2205.04766

Prasad, N., Cheng, L.-F., Chivers, C., Draugelis, M., & Engelhardt, B. (2017). *A Reinforcement Learning Approach to Weaning of Mechanical Ventilation in Intensive Care Units*.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine Learning in Medicine. *The New England Journal of Medicine*, *380*(14), 1347–1358. https://doi.org/10.1056/NEJMra1814259

Saeidi, N., Karshenas, H., & Mohammadi, H. M. (2019). Single Sample Face Recognition Using Multi cross Pattern and Learning Discriminative Binary Features. *Journal of Applied Security Research*, *14*(2), 169–190.

Sahu, M., Gupta, R., Ambasta, R. K., & Kumar, P. (2022). Chapter Three - Artificial intelligence and machine learning in precision medicine: A paradigm shift in big data analysis. In D. B. Teplow (Ed.), *Precision Medicine* (Vol. 190, Issue 1, pp. 57–100). Academic Press. https://doi.org/https://doi.org/10.1016/bs.pmbts.2022.03.002

Sanchez-Martinez, S., Camara, O., Piella, G., Cikes, M., González-Ballester, M. Á., Miron, M., Vellido, A., Gómez, E., Fraser, A. G., & Bijnens, B. (2022). Machine Learning for Clinical Decision-Making: Challenges and Opportunities in Cardiovascular Imaging. *Frontiers in Cardiovascular Medicine*, *0*, 2020. https://doi.org/10.3389/FCVM.2021.765693

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine 2019 25:1*, *25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Uddin, S., Ong, S., & Lu, H. (2022). Machine learning in project analytics: a data-driven framework and case study. *Scientific Reports*, *12*(1), 15252. https://doi.org/10.1038/s41598-022-19728-x

Vafaei, N., Delgado-Gomes, V., Agostinho, C., & Jardim-Goncalves, R. (2022). Analysis of Data Normalization in Decision-Making Process for ICU's Patients During the Pandemic. Procedia Computer Science, 214, 809–816.

Wang, H., Wang, Z., Wang, W., Xiao, Y., Zhao, Z., & Yang, K. (2020). A Note on Graph-Based Nearest Neighbor Search. In *undefined*.

Williams, K., Bilsland, E., Sparkes, A., Aubrey, W., Young, M., Soldatova, L., De Grave, K., Ramon, J., Clare, M., Sirawaraporn, W., Oliver, S., & King, R. (2015). Cheaper faster drug development validated by the repositioning of drugs against neglected tropical diseases. *Journal of The Royal Society Interface*, *12*, 20141289. https://doi.org/10.1098/rsif.2014.1289

Wong, A. N. N., He, Z., Leung, K. L., To, C. C. K., Wong, C. Y., Wong, S. C. C., Yoo, J. S., Chan, C. K. R., Chan, A. Z., Lacambra, M. D., & Yeung, M. H. Y. (2022). Current Developments of Artificial Intelligence in Digital Pathology and Its Future Clinical Applications in Gastrointestinal Cancers. *Cancers*, *14*(15). https://doi.org/10.3390/cancers14153780

Xue, Y., Du, N., Mottram, A., Seneviratne, M., & Dai, A. M. (2020). Learning to Select Best Forecast Tasks for Clinical Outcome Prediction. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, & H. Lin (Eds.), *Advances in Neural Information Processing Systems* (Vol. 33, pp. 15031–15041). Curran Associates, Inc. https://proceedings.neurips.cc/paper/2020/file/abc99d6b9938aa86d1f30f8ee0fd169f-Paper.pdf

Ying, H., Zhuang, F., Zhang, F., Liu, Y., Xu, G., Xie, X., Xiong, H., & Wu, J. (2018). Sequential recommender system based on hierarchical attention network. *IJCAI International Joint Conference on Artificial Intelligence*, *2018-July*, 3926–3932. https://doi.org/10.24963/IJCAI.2018/546

Zheng, Y., Jiang, Z., Zhang, H., Xie, F., Shi, J., & Xue, C. (2021). *Histopathology WSI Encoding based on GCNs for Scalable and Efficient Retrieval of Diagnostically Relevant Regions*. https://arxiv.org/abs/2104.07878v1

Zhou, C., Bai, J., Song, J., Liu, X., Zhao, Z., Chen, X., & Gao, J. (2018). ATRank: An attention-based user behavior modeling framework for recommendation. *32nd AAAI Conference on Artificial Intelligence, AAAI 2018*.