Proceeding Paper

Artificial Intelligence as an Emerging Tool for Cardiologists †

Łukasz Ledziński * and Grzegorz Grzesk

Department of Cardiology and Clinical Pharmacology, Faculty of Health Sciences, Collegium Medicum in Bydgoszcz, Nicolaus Copernicus University in Toruń, Ujejskiego 75, 85-168 Bydgoszcz, Poland

* Correspondence: 503345@doktorant.umk.pl

Abstract: In the world of data, there is an urgent need to find ways to extract knowledge and information for improving patient care. Artificial intelligence (AI) is an emerging tool that has the potential to provide cardiologists with new insights and knowledge. The healthcare industry has already begun the digital transformation of vast reams of data (Big Data) that are generated in routine clinical practice. AI has the potential to make a significant impact on healthcare by improving the efficiency of clinical care, providing personalized treatment, and identifying new disease biomarkers. Machine learning (ML) and deep learning (DL) are AI techniques that utilize large datasets and computational power for analysis and decision making. There are three main ML techniques: supervised learning, unsupervised learning, and reinforcement learning. Another functional AI service that has been presented is natural language processing (NLP), and it is applicable for analyzing patient documentation. In this paper, the scope of AI workflow, the most often used algorithms, and their performance metrics are explained. Explainable artificial intelligence (XAI) has a prominent potential to be a useful tool for clinicians as it provides full transparency into an AI model’s decision-making process, but few applications have been reviewed. In this paper, the challenges and limitations of AI in cardiology are discussed in terms of ethical, methodological, and legal issues. Furthermore, the successful establishment of good practices toward the right development and deployment of automated ML-based systems will ensure a regulatory framework that can strengthen patients’ trust in AI/ML-based clinical decision support systems.

Keywords: artificial intelligence; cardiology; machine learning

1. Introduction

There is a growing need for finding a new approach to handle the increasing amount of data being collected in various industries, especially in healthcare. The International Data Corporation reported a vast amount of healthcare data generated in 2013, and the amount is expected to grow rapidly in the near future. With the rate of increase, it is projected to reach a zettabyte scale soon [1]. The healthcare industry generates a large amount of information, and digital transformation is already underway, with patient information being stored in Electronic Medical Records (EMRs) [2]. The volume of healthcare data is growing rapidly, including data from healthcare professionals, wearable sensors, and applications focused on improving therapy adherence. Health data are complex and characterized by high volume, velocity, veracity, and variety—making them difficult to handle with simple methods. Artificial intelligence (AI) has the potential to revolutionize the healthcare system, having already made a significant impact on molecular chemistry and drug discovery [3]. AI includes systems that can learn and make decisions, and machine learning (ML) is a subpart of AI that can develop systems that learn from retrospective data and create classification, clustering, or regression models. Machine learning is already being used in screening and diagnostic medical models and can help interpret results and make more personalized
clinical decisions [4–6]. In comparison to traditional statistics, ML has a bigger scope on developing automated clinical decision systems than scoring systems [7–9].

2. What Is AI?

The process of obtaining a machine learning model starts by defining the objective and gathering data from various sources to create a large dataset. The data are then prepared by removing any missing values or unnecessary information, followed by exploring and splitting the data. The next step is to train the model using this dataset and then tune its hyperparameters, followed by testing and validation. The first step in the machine learning process is to identify the problem and determine what type of input data is available, which features are targeted, and what type of problem it is. Gathering data from EMRs and curating them can be time-consuming and challenging. Exploratory data analysis is important to understand any correlations, associations, patterns, and trends in the data.

It is important to understand the three main types of machine learning (ML) before starting to build models: supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, a model is generated using labeled data with a known outcome of interest for the training set. The model adjusts the weights until it is fitted for classification or regression. Classification analysis focuses on assigning a class to each object, such as determining if a patient has a disease. The output of a classification model is typically nominal. On the other hand, regression models typically have continuous output that can be used, for example, to estimate the likelihood of a disease. Cross-validation must be performed after the learning process to ensure that the model is not underfitting or overfitting. An underfitting model will not accurately predict or classify unseen data and training data, while an overfitting model is ideal for training data but cannot perform accurately on unknown data [10]. The workflow for implementing artificial intelligence is shown in Figure 1.

![Figure 1. AI project—from idea to model implementation.](image-url)

The unsupervised learning process involves utilizing unlabeled data to uncover patterns and grouping similar data points together through clustering or anomaly detection, which is commonly applied in the field of medicine. Reinforcement learning, on the other hand, involves training an algorithm on dynamic data and constantly adjusting it to improve the results by trial and error, with the goal of achieving the maximum reward.

Comparison of various models needs to be based on metrics that identify how good a model is in solving a problem. Performance metrics are different for classification and regression. Principal metrics and their descriptions are listed in Table 1.
Table 1. Performance metrics.

<table>
<thead>
<tr>
<th>Performance Metrics</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Classification</strong></td>
<td></td>
</tr>
<tr>
<td>Precision (PPV)</td>
<td>The fraction of objects correctly classified as positive among all positive classifications</td>
</tr>
<tr>
<td>Sensitivity (TPR)/Recall</td>
<td>The fraction of objects correctly classified as positive that are actually positive</td>
</tr>
<tr>
<td>Accuracy</td>
<td>The fraction of objects correctly classified as positive and the fraction of objects correctly classified as negative</td>
</tr>
<tr>
<td>F1 score</td>
<td>The harmonic mean of precision and recall</td>
</tr>
<tr>
<td>Specificity (TNR)</td>
<td>The fraction of objects correctly classified as negative that are actually negative</td>
</tr>
<tr>
<td>Receiver operating characteristic (ROC) curve</td>
<td>The curve between recall (Y-axis) and specificity (X-axis), with the false positive rate = 1</td>
</tr>
<tr>
<td>Area under the curve</td>
<td>The AUC evaluates the overall quality of the model</td>
</tr>
<tr>
<td><strong>Regression</strong></td>
<td></td>
</tr>
<tr>
<td>Mean absolute error (MAE)</td>
<td>The mean of the absolute difference between the actual and predicted values in a dataset</td>
</tr>
<tr>
<td>Mean squared error (MSE)</td>
<td>The mean squared error between the predicted and actual values in a dataset</td>
</tr>
<tr>
<td>Root mean squared error (RMSE)</td>
<td>Square root of MSE</td>
</tr>
<tr>
<td>Coefficient of determination ($R^2$)</td>
<td>The proportion of variance explained by the model</td>
</tr>
</tbody>
</table>

3. Natural Language Processing

Natural language processing (NLP) is an important feature of artificial intelligence that is widely used in medicine. NLP enables the automatic analysis and representation of human language using computational methods. The versatility of NLP in medicine ranges from text classification and information extraction to search engines, enabling effective use of clinical notes. Another aspect of NLP is the use of speech recognition and question answering to develop chatbots for patients [11]. Combining NLP with ML methods provides an opportunity to extract valuable insights from large amounts of unstructured text in EMR data, which can be used to create a clinical decision-making support system, diagnostic tool, or recommendation system based on a search engine for scientific papers.

4. AI Applications in Cardiology

4.1. Supervised Learning

Supervised learning has proven to be useful in cardiovascular medicine for various applications [12,13]. Artificial neural networks (ANN) have been used for ECG interpretation to identify life-threatening arrhythmias [14]. Another significant research has been performed by training a subtype of ANN—convolutional neural network (CNN)—on paired 12-lead electrocardiography (ECG) and echocardiography data to identify patients with ventricular dysfunction, achieving an accuracy of 85.7% [15]. Echocardiography data have also been used to train another CNN model to detect regional wall motion abnormalities, achieving a result similar to that of cardiologists [16]. In addition, decision trees have been used for the classification of patients with different types of atrial fibrillation [17]. A support vector machine (SVM) model was developed to detect the severity of mitral regurgitation by analyzing 2D echocardiography videos with high sensitivity (99.38%), specificity (99.63%), and accuracy (99.45%) [18]. A study comparing AI applications for diagnosing acute coronary syndrome found that SVM outperformed ANN, Naïve Bayesian, and logistic regression [19]. LogitBoost, a variation of decision trees, was used by a group of researchers to predict mortality in patients with coronary artery disease, outperforming
the standard Framingham risk score [20]. Machine learning algorithms have been shown to be better than conventional statistical models for predicting readmission and mortality in heart failure patients [21]. AI can also be used to identify predictors of acute coronary syndrome events or to diagnose pulmonary arterial hypertension [22–24].

4.2. Unsupervised Learning

Unsupervised learning is mainly employed for tasks such as grouping, clustering, and dimensionality reduction. In one study, the researchers utilized hierarchical clustering to predict the effect of β-blocker therapy on heart failure patients based on their left ventricular ejection fraction [25]. In another study, unsupervised learning algorithms were applied to combine standard clinical parameters and echocardiographic images to create a more clinically interpretable classification of a diverse group of heart failure patients and to identify those who would be more likely to respond to treatment [26]. Treatment personalization based on AI/ML models could be a chance for extending the lifespan of patients. It would be crucial for rare and lethal diseases. Although unsupervised learning is less widely used than supervised learning, it has found applications in automating the analysis of EMR or genetic data [27].

4.3. Reinforcement Learning

In cardiology, reinforcement learning has had limited applications, but it has potential for personalizing therapy based on a patient’s characteristics. In one study, the SARSA (state–action–reward–state–action) reinforcement learning algorithm was used to determine the dosage of dofetilide based on the negative consequences of unsuccessful initiation [28]. Another study developed a high-performing and robust method for detecting anatomical landmarks in real time [29].

4.4. Natural Language Processing

NLP is primarily a method for making unstructured text data readable for machines. Although there are numerous applications of NLP outside of healthcare, such as chatbots and search engines, there is potential for more research in the field due to the improvement seen in algorithm performance from incorporating unstructured data. NLP can be used as an efficient tool for testing peripheral arterial disease based on clinical narrative notes and has been shown to outperform billing code algorithms, with a sensitivity of 91.2%, specificity of 92.5%, and accuracy of 91.8% [30]. Additionally, researchers have found that merging clinical and demographic features with narrative data from EMRs can greatly improve the efficiency of a model when using NLP [31]. Using NLP in medical studies may also reduce misclassifications by providing additional information found only in EMRs.

4.5. Explainable AI

Increasing concerns about the safety, responsibility, and reliability of AI/ML-based systems developed for the healthcare industry leads to a heightened focus on explainable AI (XAI). This is due to the primary usage of supervised learning and employment of unsupervised learning for tasks such as reducing dimensions, clustering, and identifying relationships between observations. Oncologists highlight errors made by IBM Watson in oncology. In previous research, application created for suggesting cancer treatments was trained on unrealistic and biased data that were not labeled based on guidelines and evidence-based medicine, which caused many incorrect recommendations being made in the treatment of patients [32–34]. It is important to realize that medical mistakes made by human doctors affect a limited number of patients, while a flawed AI model could impact the majority of people in a community. XAI provides insights into how AI makes decisions, including the strengths and weaknesses of the program used, the reasoning behind a specific decision, and the specific criteria used to make decisions. To address the challenge of explainability, new technologies, such as GradCAM (Gradient-weighted Class Activation Mapping) and SHAP (SHapley Additive exPlanations) algorithms, are
being developed [35–37]. An XAI algorithm combined with gradient tree boosting was applied in research to develop a model for predicting pulmonary hypertension based on EMRs [38].

5. Legal, Ethical, and Methodological Issues

There is a need for official international requirements for the deployment of AI models in medicine. There are already some guidelines, but many questions remain unresolved. The use of AI in medicine can help overcome some limitations, but it also presents challenges related to the algorithm/model and data used. The most important algorithm-related problems are standardization, reproducibility, explainability, and legal/ethical responsibility for patient outcomes. The World Health Organization has written ethical guidance for AI in healthcare [39].

One of the main challenges related to data in the deployment of AI in medicine is the acquisition and safety of data. The transfer of patient identifiable EHRs without permission, as seen in the case of the Royal Free London NHS Foundation Trust and Google Deep Mind, has raised public trust concerns [40]. European regulations on data protection and privacy are well established with the General Data Protection Regulation [41]. Another challenge is data quality, and it is crucial to develop international EHR registries that are robust, transparent, trustworthy, and verifiable. There also needs to be a solid and frequently checked security system before full AI implementation to prevent hacking and ensure patient safety. Recent research has revealed vulnerabilities in AI healthcare systems and the FDA has issued warnings about the vulnerability of medical devices to cyberattacks [42–44].

6. Discussion

The use of AI in medicine is crucial for improving patient outcomes and reducing physician burnout, as well as facilitating the clinical decision-making process, especially in the treatment of rare and lethal diseases. AI is seen as a complement to traditional forms of care, with the potential to improve diagnostic accuracy and provide personalized therapy. However, it also has some disadvantages and pitfalls that must be considered, and it is important to be aware of these when using AI-based technology in medicine [45].

Author Contributions: Conceptualization, Ł.L. and G.G.; writing—original draft preparation, Ł.L.; writing—review and editing, G.G.; visualization, Ł.L.; supervision, G.G.; funding acquisition, G.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.