

Review

# What Is Machine Learning, Artificial Neural Networks and Deep Learning?—Examples of Practical Applications in Medicine

Jakub Kufel <sup>1,\*</sup>, Katarzyna Bargiel-Łączek <sup>2</sup>, Szymon Kocot <sup>3</sup>, Maciej Koźlik <sup>4</sup>, Wiktoria Bartnikowska <sup>2</sup>, Michał Janik <sup>5</sup>, Łukasz Czogalik <sup>5</sup>, Piotr Dudek <sup>5</sup>, Mikołaj Magiera <sup>5</sup>, Anna Lis <sup>6</sup>, Iga Paszkiewicz <sup>5</sup>, Zbigniew Nawrat <sup>1</sup>, Maciej Cebula <sup>7</sup> and Katarzyna Gruszczynska <sup>8</sup>

- <sup>1</sup> Department of Biophysics, Faculty of Medical Sciences in Zabrze, Medical University of Silesia, 41-808 Zabrze, Poland; znawrat@sum.edu.pl or nawrat@frk.pl
  - <sup>2</sup> Paediatric Radiology Students' Scientific Association at the Division of Diagnostic Imaging, Department of Radiology and Nuclear Medicine, Faculty of Medical Science in Katowice, Medical University of Silesia, 40-752 Katowice, Poland; katarzyna.bargiel@op.pl (K.B.-Ł.); wiktoriabart96@gmail.com (W.B.)
  - <sup>3</sup> Bright Coders' Factory, Technologiczna 2, 45-839 Opole, Poland
  - <sup>4</sup> Division of Cardiology and Structural Heart Disease, Medical University of Silesia, 40-635 Katowice, Poland; kozlik.maciej@gmail.com
  - <sup>5</sup> Student Scientific Association Named after Professor Zbigniew Religa at the Department of Biophysics, Faculty of Medical Sciences in Zabrze, Medical University of Silesia, Jordana 19, 41-808 Zabrze, Poland; michal.janik0801@gmail.com (M.J.); lukczog@gmail.com (Ł.C.); piotrek233@gmail.com (P.D.); mikaczu7422@gmail.com (M.M.); igapaszkiewicz.ip@gmail.com (I.P.)
  - <sup>6</sup> Cardiology Students' Scientific Association at the III Department of Cardiology, Faculty of Medical Sciences in Katowice, Medical University of Silesia, 40-635 Katowice, Poland; lis.anna9898@gmail.com
  - <sup>7</sup> Individual Specialist Medical Practice Maciej Cebula, 40-754 Katowice, Poland; maciejmichalcebula@gmail.com
  - <sup>8</sup> Department of Radiodiagnostics, Invasive Radiology and Nuclear Medicine, Department of Radiology and Nuclear Medicine, School of Medicine in Katowice, Medical University of Silesia, Medyków 14, 40-752 Katowice, Poland; kgruszczynska@sum.edu.pl
- \* Correspondence: jakubkufel92@gmail.com



**Citation:** Kufel, J.; Bargiel-Łączek, K.; Kocot, S.; Koźlik, M.; Bartnikowska, W.; Janik, M.; Czogalik, Ł.; Dudek, P.; Magiera, M.; Lis, A.; et al. What Is Machine Learning, Artificial Neural Networks and Deep Learning?—Examples of Practical Applications in Medicine. *Diagnostics* **2023**, *13*, 2582. <https://doi.org/10.3390/diagnostics13152582>

Academic Editors: Dechang Chen, Qaisar Abbass and Abdul Rauf Baig

Received: 24 June 2023

Revised: 19 July 2023

Accepted: 1 August 2023

Published: 3 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Machine learning (ML), artificial neural networks (ANNs), and deep learning (DL) are all topics that fall under the heading of artificial intelligence (AI) and have gained popularity in recent years. ML involves the application of algorithms to automate decision-making processes using models that have not been manually programmed but have been trained on data. ANNs that are a part of ML aim to simulate the structure and function of the human brain. DL, on the other hand, uses multiple layers of interconnected neurons. This enables the processing and analysis of large and complex databases. In medicine, these techniques are being introduced to improve the speed and efficiency of disease diagnosis and treatment. Each of the AI techniques presented in the paper is supported with an example of a possible medical application. Given the rapid development of technology, the use of AI in medicine shows promising results in the context of patient care. It is particularly important to keep a close eye on this issue and conduct further research in order to fully explore the potential of ML, ANNs, and DL, and bring further applications into clinical use in the future.

**Keywords:** AI; artificial intelligence; medicine; AI in medicine

## 1. Introduction

Artificial intelligence (AI) has gained significant attention over the past decade, with machine learning (ML) and deep learning (DL) being popular topics [1,2]. While these terms are related, they have distinct meanings and cannot be used interchangeably. This article aimed to provide clinicians with accessible information on the AI methods used

in medicine without relying heavily on technical jargon. Various definitions of AI exist, with some focusing on intelligent systems making decisions to achieve goals, while others emphasise machines mimicking cognitive functions, like learning and problem solving. ML and DL are specific methods within the broader field of AI, contributing to the automation of intellectual tasks performed by humans. This article will also cover FDA-approved AI solutions in medicine, discussing the machine learning methods that have been employed, the type of data integrated, their overall performances, as well as their advantages and disadvantages. The content of this study will be presented in a manner understandable to those without technical expertise in the field.

## 2. Glossary

- **Model architecture**—describes how the data is processed, transmitted, and analysed within the machine learning algorithm, which influences its efficiency and effectiveness in solving problems.
- **Data exploration**—the process of analysing and summarising a large dataset to gain insight into the relationships and patterns that exist within the data.
- **Binary classification**—a type of classification in which the aim is to assign one of two possible classes (labels) to an object: positive or negative, true or false, etc.
- **Logistic function**—a sigmoidal mathematical function that transforms values from minus infinity into plus infinity to the range (0, 1), allowing not only non-linearity, but also probability, e.g., binary classification.
- **Input variables**—also known as independent variables, explanatory variables, predictor variables, etc., and are variables that are used to describe or explain the behaviour, trends, or decisions of the target variables.
- **Target variables**—also known as dependent variables, outcome variables, etc., and are variables that are studied or predicted in statistical analysis and machine learning. Target variables are dependent on and are described using explanatory variables.
- **Bayes theory**—used to calculate the probability of an event, having prior information about that event.
- **Sentiment analysis**—the process of automatically determining emotions, opinions, and moods expressed in the text. This can be in the form of product reviews, comments on online forums, tweets on Twitter, or other forms of textual communication. The purpose of sentiment analysis is to gain an automatic understanding of whether a text is positive, negative, or neutral.
- **Training**—the process of formatting a model to interpret the data to perform a specific task with a specific accuracy. In this case, it is the determination of the hyperplane.
- **Hyperplane**—a set having  $n - 1$  dimension, relative to the  $n$ -dimensional space in which it is contained (for  $n =$  two-dimensional space it has one dimension (point); for  $n =$  three-dimensional space it has two dimensions (line)).
- **Training objects**—a set of objects used to determine the hyperplane with the model.
- **Support vectors**—at least two objects at the shortest distance from the hyperplane belonging to two classes.
- **Class**—a group, described on numerical ranges, to which an object can be assigned—i.e., classified.
- **Cluster**—a hyperplane-limited space in a data system in which the presence of an object determines the class assignment.
- **Neuron**—the basic element of a neural network, which connects to other neurons through transmitting data to each other.
- **Weight**—the characteristic that the network designer provides to the connections between the neurons to achieve the desired results.
- **Recursion**—referring a function to the same function using the network being trained.
- **Layer**—a portion of the total network architecture that is differentiated from the other parts due to having a distinctive function.

- **Activation function**—takes the input from one or more neurons and maps it to an output value, which is then passed onto the next layer of neurons in the network.
- **Hidden layer**—in an artificial neural network, this is defined as the layer between the input and output layers, where the result of their action cannot be directly observed.
- **Input layer**—layer where the data are collected and passed onto the next layer.
- **Output layer**—layer which gathers the conclusions.
- **Backpropagation**—sending signals in the reverse order to calculate the error associated with each individual neuron in an effort to improve the model.
- **Cost function**—a function that represents the errors occurring in the training process in the form of a number. It is used for subsequent optimisation.
- **Receptive field**—a section of the image that is individually analysed using the filter.
- **Filter**—a set of numbers that are used to perform computational operations on the input data on splices. It is used to extract features (e.g., the presence of edges or curvature).
- **Convolution**—integral of the product of the two functions after one is reflected about the y-axis and shifted.
- **Pooling**—reducing the amount of data representing a given area of the image.
- **Matrix**—a mathematical concept; a set of numbers which is used, among other things, to recalculate the data obtained from neurons.
- **Skip connections**—a technique used in neural networks that allows information to be passed from one layer of the network to another, while skipping intermediate layers.

### 3. Data and Its Relevance to Neural Networks

Data are essential for any data analysis method. It consists of facts or information in different forms, like sound, text, values, and images. In AI, data are used to train models and enable them to make decisions and perform tasks, like classification or text generation. The proper interpretation of data is crucial for training neural network models. Data can be structured (easily organised) or unstructured (difficult to classify). Structured data, such as phone numbers or financial data, can be processed and analysed effectively. Unstructured data are those that are not easily classified and organised; examples include images and sound files [3].

The input data mentioned above needs to be normalised for the successful learning process of a neural network. Meanwhile, the output data from the neural network are used for predictions or solutions.

To determine the training dataset size for a machine learning model, several parameters are required to be defined, which include the complexity of the model, the complexity of the learning algorithm, the need for labelling, the definition of an acceptable margin of error, and the diversity of the inputs used [4]. In short, the input data must be prepared in terms of the task and the expected results.

Data processing can be approached in two ways: focusing on the model or focusing on the data. The first approach aims to enhance the performance of the machine learning (ML) model by selecting an efficient model architecture and learning process. It involves improving the model's code without changing the amount of data. This method is gaining increasing levels of popularity as it eliminates the need for extensive data collection. The second strategy emphasises working on the data itself, and involves modifying and enhancing the datasets to improve the accuracy of the decisions being made [5].

### 4. Machine Learning Models

Machine learning is a subset of AI that involves building computer models that are capable of learning and making independent predictions or decisions based on the provided data. These models continually improve their accuracy through learned data.

The author of the name machine learning, Arthur Samuel, used this phrase in 1959 to describe “the ability of computers to learn without programming new skills directly” [6]. Using the available datasets, a machine learning algorithm, supported with a mathematical

model, generates predictions or specific decisions. The main types of machine learning are supervised and unsupervised learning.

#### 4.1. Supervised Learning

Supervised machine learning assumes that the model has been trained on a similar dataset to the problem at hand, consisting of the input data and the corresponding output data. Once the model learns the relationship between the input and the output, it can classify new unknown datasets and make predictions or decisions based on them. This type of learning is divided into two methods: classification and regression.

In supervised machine learning, e.g., a photograph that the algorithm classifies as either a cat or a dog, representing a two-class classification problem, the solutions are typically binary, in the form of either yes or no. Another example is handwriting recognition, where the software matches the handwritten characters (output data) to their corresponding printed counterparts (classes).

Regression is a fundamental type of supervised learning that predicts continuous values using input data. For instance, in healthcare, regression can be applied to forecast the medical costs. The input data would include drug prices, the required medical equipment, and staff expenses, while the output would be the total treatment cost. Through training these models with the input and output data, predictions for the total treatment cost can be made for new inputs.

#### 4.2. Unsupervised Learning

Unsupervised ML differs from supervised ML in its use of unannotated data, which has not been previously labelled by humans nor algorithms. The model learns from input data without expected values, and the available dataset does not provide answers to the given task. Instead of labelling or predicting outputs, this algorithm focuses on grouping the data based on their characteristics. The goal is to teach the machine to detect patterns and group the data without a single correct answer. There are two types of unsupervised learning: clustering and association.

Clustering involves grouping the data based on their similarities and differences. For example, animals can be divided into groups based on their visual features determined using the model.

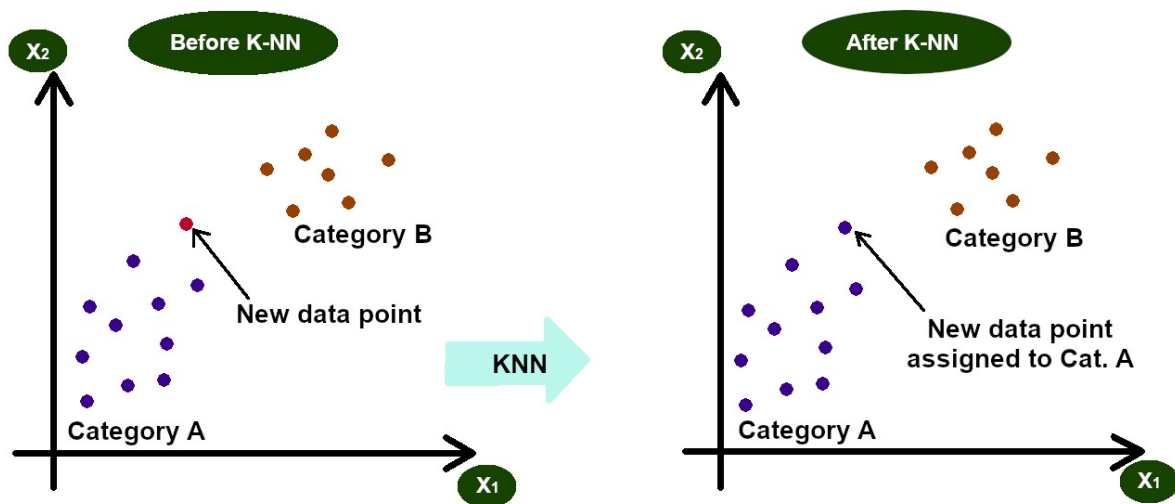
Association is a method of analysing the relationships between data in a dataset. For instance, the algorithm can pair people buying mattresses for pressure sores with those ordering products to aid in the healing of pressure sores. This method is commonly used in marketing strategies.

### 5. Classical Methods of Machine Learning

#### 5.1. *k*-Nearest Neighbour Algorithms

The *k*-nearest neighbour (kNN) method is a popular method used in data mining and statistics. The kNN method is a type of algorithm that predicts the correct class of the test data by calculating the distance between the test data and all the training points (Figure 1). It then shows the number of *k* (training) points that are close to the test data. In the case of regression, the obtained value is the average of the selected training points, “*k*” [7].

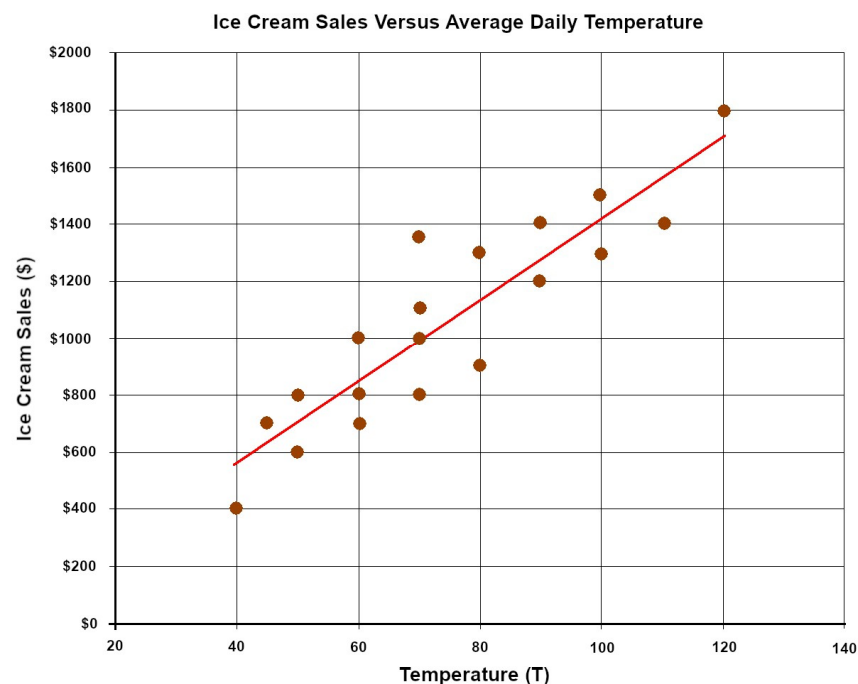
Hamed et al. adapted the kNN algorithm for handling incomplete COVID-19 datasets. They implemented a novel variant called kNNV and tested it on an Italian dataset from Italian Society of Medical and Interventional Radiology (SIRM). Their results demonstrated an effective and accurate COVID-19 case classification (with an accuracy of 0.72–1, and a precision of 0.7–1) [8]. Bellino et al. implemented the kNN algorithm for automatic classification in the operating theatre, aiding neurophysiologists and neurosurgeons in deep brain stimulation electrode fixation for Parkinson’s disease patients. This system, when refined, will reduce surgery time and improve effectiveness [9].



**Figure 1.** k-nearest neighbour (kNN) example p—new data before and after kNN.

### 5.2. Linear Regression Algorithms

Linear regression predicts the value of a dependent variable from an independent variable (an example of which has been displayed in Figure 2). It generates a simple, interpretable formula for predictions and is widely used in Python and Excel, as well as across a range of fields, like science, biology, business, and behavioural science [10].



**Figure 2.** Linear regression example—ice cream sales versus average daily temperature—individual values on subsequent days are represented by brown circles. The red line stands for the linear regression plot created from this data.

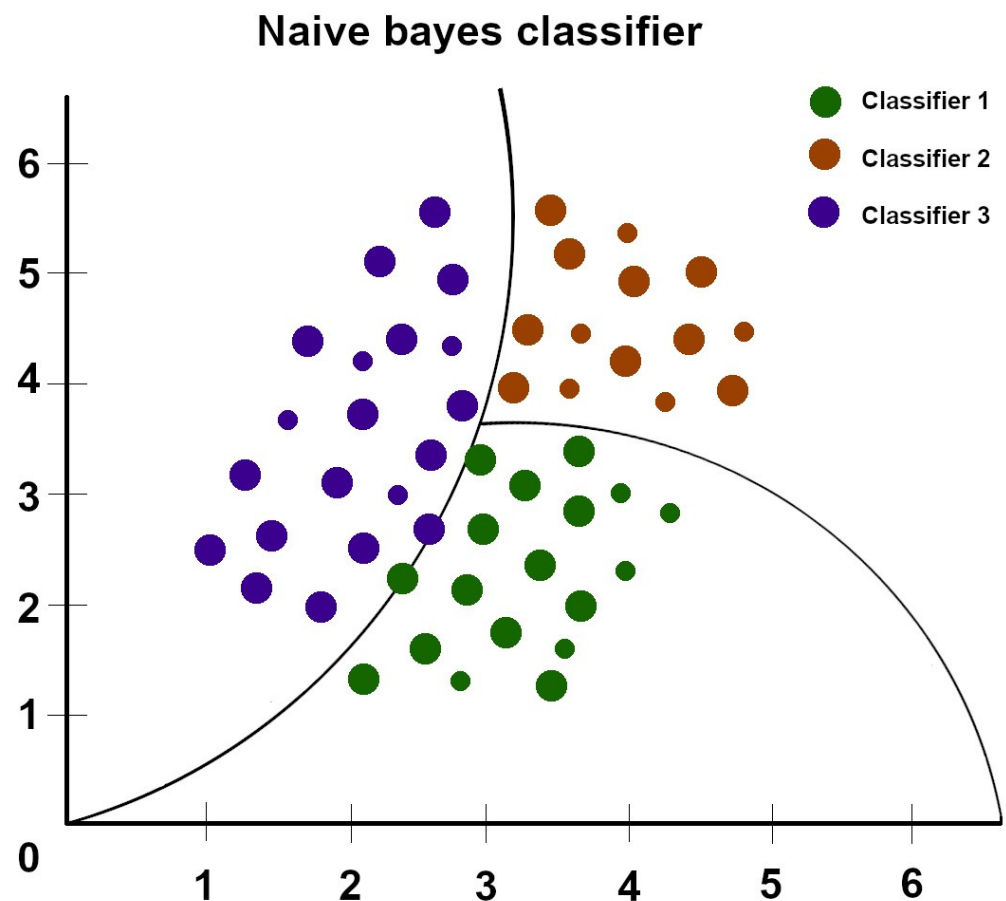
An example of the use of these linear regression algorithms is the study presented by Garcia et al. on estimating the carotid-to-femoral pulse wave velocity (cf-PWV) using multiple linear regression (MLR). This model used blood pressure and/or photoplethysmography. Their results suggest that ML (specifically MLR) combined with a semi-classical signal analysis method could be a valuable automated tool for efficient cf-PWV assessments in the future [11].

### 5.3. Logistic Regression Algorithms

Logistic regression predicts the probability of an object belonging to one of two classes. It uses a logistic function to transform predicted values between zero and one. This algorithm establishes the relationship between the input variables and the target variable. Logistic regression has found its application in market research, medical analysis, and banking to understand variable relationships and predict their likely outcomes, such as illness or customer purchases. It helps in answering the questions concerning the likelihood of events or group membership based on the characteristics of the data, like age and education [12–15].

### 5.4. Naïve Bayes Classifier Algorithms

Sentiment analysis automatically identifies emotions, opinions, and moods in the text, such as product reviews, forum comments, and “tweets”. It determines whether the text is positive, negative, or neutral. The Naïve Bayes classifier assumes feature independence and quickly categorises information based on this assumption (Figure 3). It is widely used in spam filtering, text classification, and sentiment analysis [16–18].

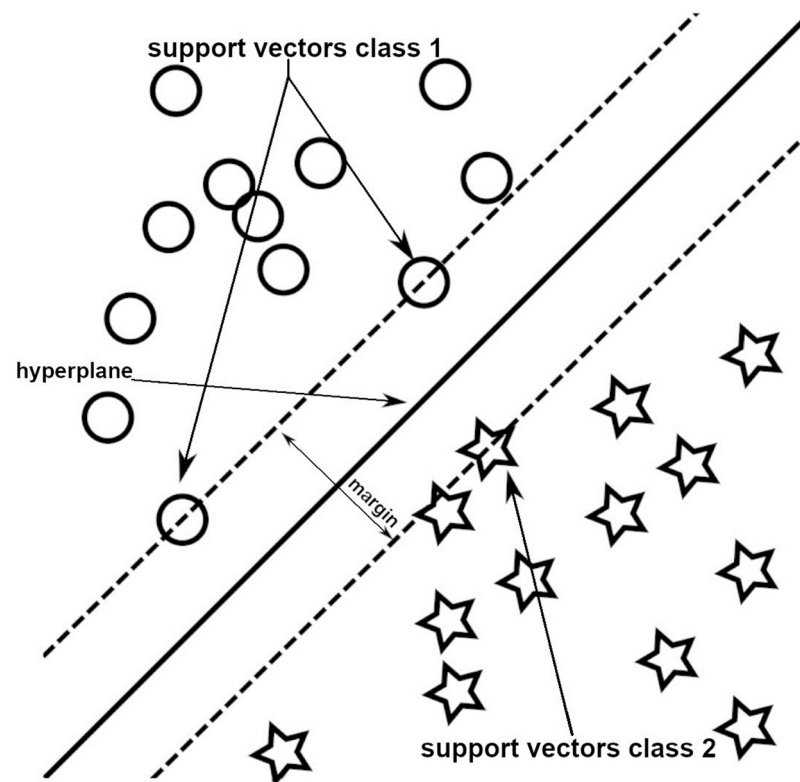


**Figure 3.** A function graphically depicting the performance of a Naive Bayes classification algorithm.

### 5.5. Support Vector Machines (SVMs)

Support vector machines (SVMs) are classed as another type of machine learning model that can be used to classify objects. SVMs involve finding the hyperplane that best separates objects from two different classes (Figure 4). However, in order to allocate these objects to more than one class, separate binary classification training is required [19].





**Figure 4.** A simplified example of the support vectors and samples of two classes.

In an  $n$ -dimensional data system ( $n \geq 1$ ), objects are mapped based on their properties. A machine determines a separating hyperplane that maximises the margin between two classes of training objects [20]. These objects are referred to as support vectors [21]. The hyperplane can be a point, line, or plane depending on the object's dimensions. Its purpose is to define the class membership by assessing which side of the boundary an object falls on. In some cases, an object may possess the properties of another class, leading to the formation of a soft margin, where objects can be located outside the defined cluster without affecting the hyperplane position. If a single hyperplane cannot separate two data clusters, a kernel function can be used to transform the data and achieve a linear separation of the classes [22].

SVMs are valued for their ability to handle non-linear data and high-dimensional classification tasks, making it suitable for processing large datasets [23]. It excels at recognising intricate patterns in complex datasets and finds applications across diverse areas, like handwriting recognition, credit card fraud detection, and facial recognition [24].

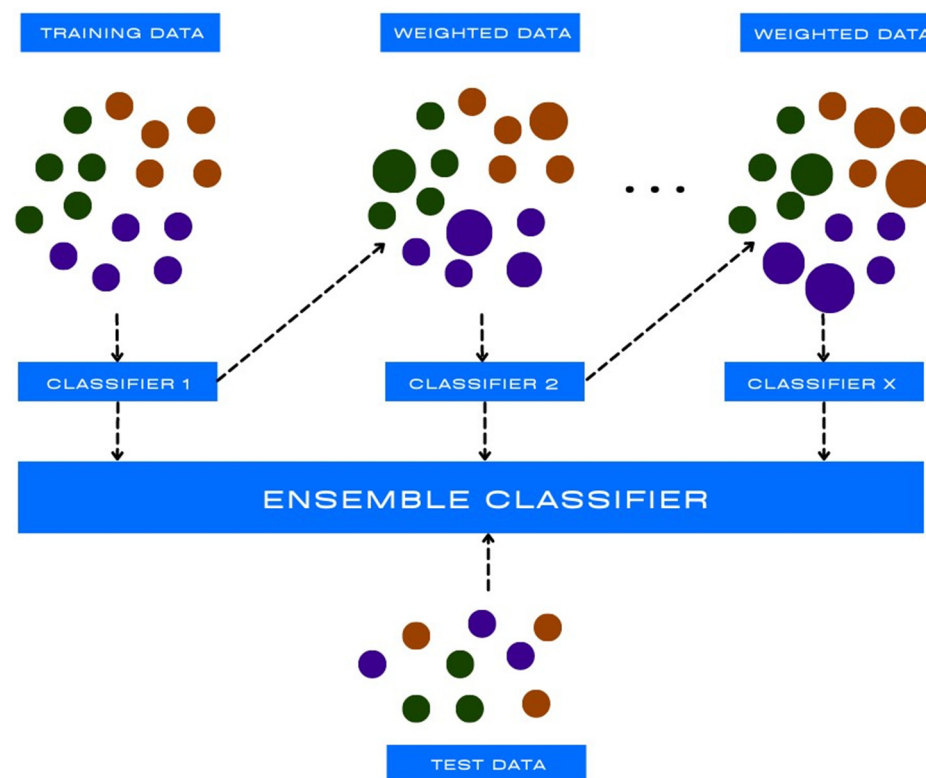
In a study by Zhang et al., SVMs were used to diagnose diabetes using tongue photographs. Tongue images were captured from 296 diabetic patients and 531 non-diabetic patients using a TDA-1 camera, respectively. These researchers focused on colour and texture analysis of the tongue's body and surface. SVM models were trained using this data, resulting in an algorithm that achieved a superior efficiency of 79.72% in diabetes detection from tongue images compared to existing methods [25].

#### 5.6. AdaBoost

Adaboost, short for adaptive boosting, is a machine learning algorithm used for binary classification and regression tasks. It is a relatively new non-linear machine learning algorithm [26]. AdaBoost is a boosting technique that iteratively trains multiple weak models on different subsets of the training data and assigns higher weights to the misclassified instances in each iteration (as shown in Figure 5). In subsequent iterations, the algorithm focuses more on the misclassified samples, allowing the weak models to learn from their

mistakes and improve on their performance, following which they are combined to form a single strong classifier [27].

Adaboost is widely used in computer-aided diagnosis (CAD) and can support medical practitioners to make critical decisions regarding their patients' disease conditions, such as Alzheimer's disease, diabetes, hypertension, and various cancers [28]. Furthermore, it can help to improve the accuracy in classifying diseases, predicting patient outcomes, and detecting abnormalities in medical images [29,30].



**Figure 5.** Diagram of the AdaBoost algorithm—different size circles stand for samples with more associated weights, various colors of the circles stand for different subsets of data [31].

### 5.7. XGBoost

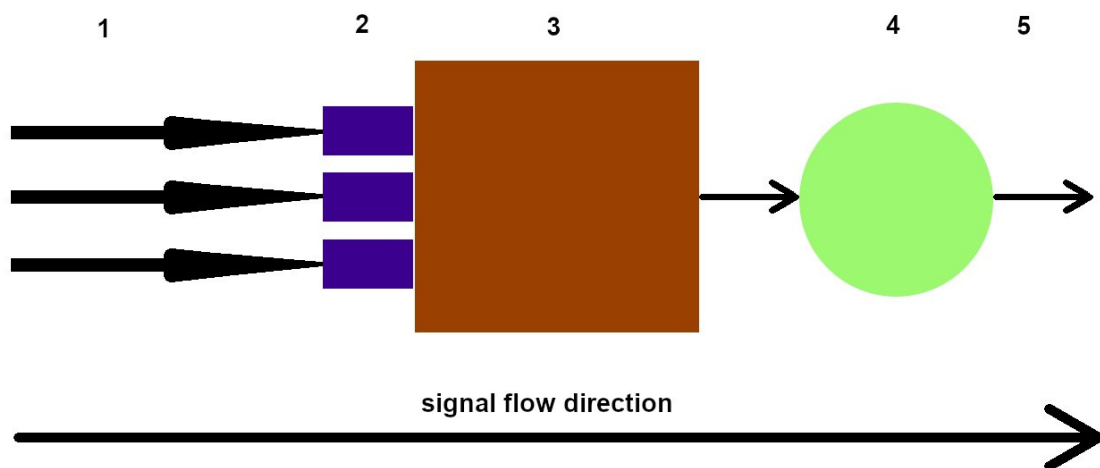
XGBoost, which stands for extreme gradient boosting, is a powerful machine learning algorithm that has gained significant popularity across various domains, including in medicine. It is an ensemble learning method that aggregates the predictions of many individually trained weak decision trees to create a more accurate and more powerful model [32].

In medicine, XGBoost has been successfully applied to a wide range of tasks, such as disease diagnosis, prognosis, treatment selection, and patient outcome prediction [33–35]. One of the key strengths of the XGBoost algorithm is its ability to handle diverse types of medical data, including structured data (e.g., patient demographics and laboratory results) and unstructured data (e.g., medical images and clinical notes).

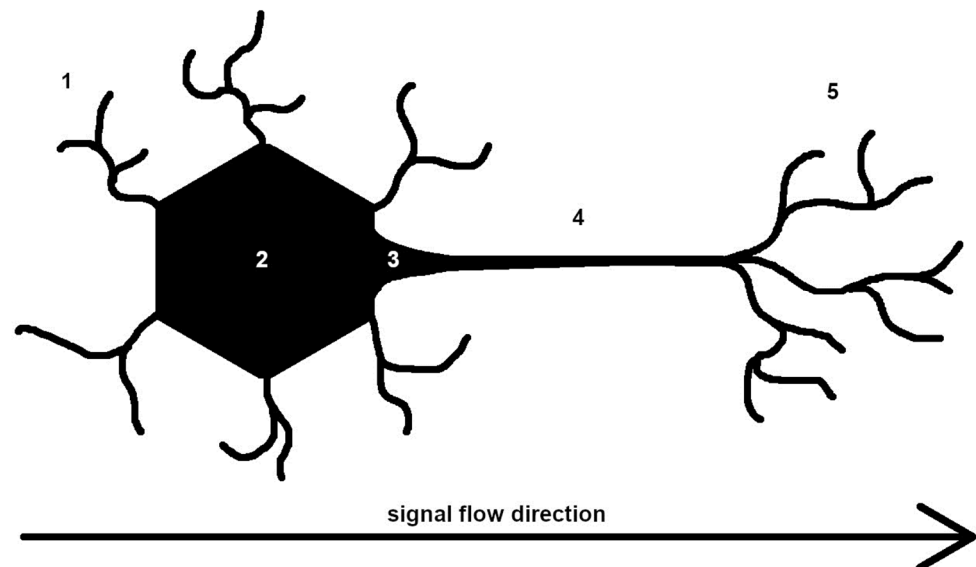
## 6. Neural Networks

Artificial neural networks resemble the human brain (Figure 6) and comprise multiple perceptrons or 'neurons' that process and transmit information. In understanding their functioning (Figure 7), we begin with inputting data, such as images, text, or sound. This data traverses the network, processed using successive layers of perceptrons until reaching the output. Each layer contains multiple neurons that process the input data.





**Figure 6.** A simplified diagram of a mathematical neuron. 1—signal inputs, 2—scales, 3—adder, 4—activator (activation function), and 5—signal output, respectively.



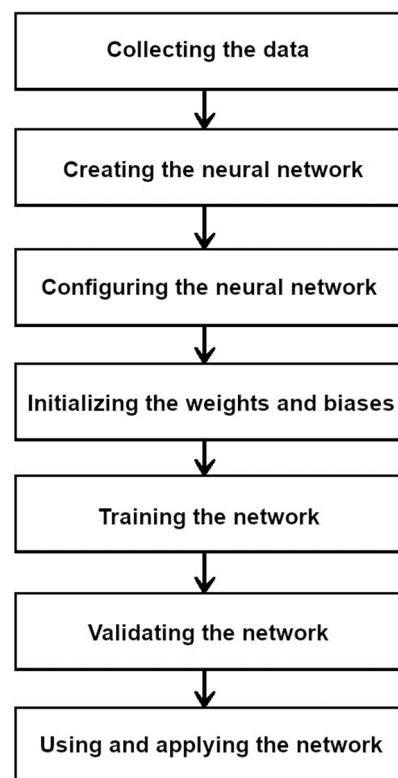
**Figure 7.** A simplified diagram of a human neuron. 1—dendrites, signal input site, 2—nucleus of the neuron, 3—zone of initiation (where the action potential of the neuron is formed), 4—axon, and 5—axon terminals (which form connections with other cells, and are the sites of signal output), respectively.

Neural networks are widely used in image recognition, natural language processing, speech recognition, and stock price prediction. They consist of interconnected neurons and come in various types, as listed below:

1. Perceptron networks: simplest neural networks with an input and output layer composed of perceptrons. Perceptrons assign a value of one or zero based on the activation threshold, dividing the set into two.
2. Layered networks (feed forward): multiple layers of interconnected neurons where the outputs of the previous layer neurons serve as the inputs for the next layer. The neurons of each successive layer always have a +1 input from the previous layer. Enables the classification of non-binary sets, and are used in image, text, and speech recognition.
3. Recurrent networks: neural networks with feedback loops where the output signals feed back into the input neurons. Can generate sequences of phenomena and signals until the output stabilises. Used for sentiment analysis and text generation.
4. Convolutional networks, also known as braided networks, are described in the next paragraph.

5. Gated recurrent unit (GRU) and long short-term memory (LSTM) networks: perform recursive tasks with the output dependent on previous calculations. They have network memory, allowing them to remember data states across different time steps. These networks have longer training times and are applied in time series analysis (e.g., stock prices), autonomous car trajectory prediction, text-to-speech conversion, and language translation.

In order to train the perceptrons, weights are adjusted to minimise the difference between the output and the expected signal. The network also learns through the greatest gradient decline method, adjusting the step lengths in the opposite direction. If the target value at a new point surpasses the starting point, the steps are reduced until the desired value has been achieved (Figure 8).



**Figure 8.** Simplified diagram of the neural network operation.

Backpropagation is another type of ML method that calculates the error for neurons in the last layer and propagates it backwards to the earlier layers. This efficient algorithm has been widely used in research. This network can be tested with new data to assess its performance in recognizing previously unseen information.

#### *Vision Transformers*

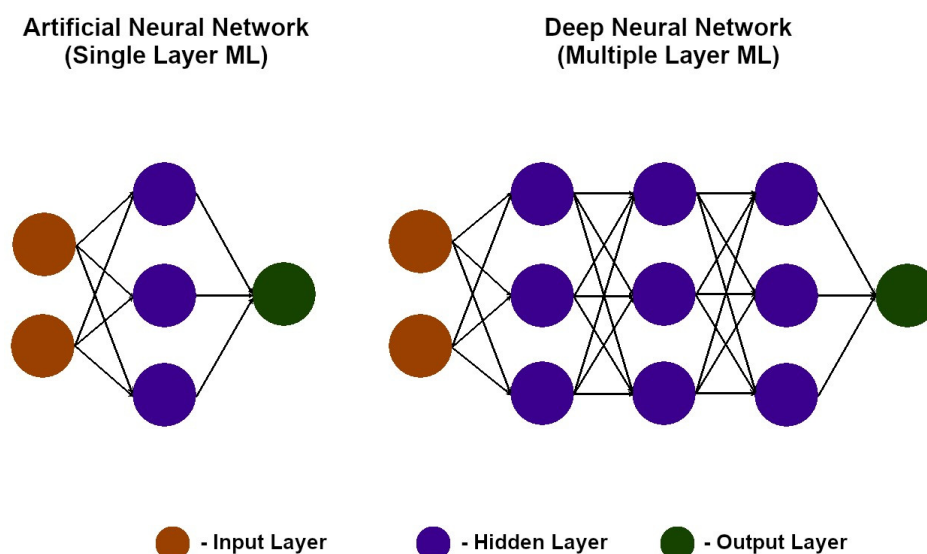
As transformers have become more and more successful in solving NLP tasks, they can be used as powerful alternatives to CNNs [36]. Vision transformers have the input image split into rectangular patches, and with the position index combined, they are processed with the standard transformer encoder block without requiring any convolutional layers and reduce the risk of adversarial attacks as a result [37].

## **7. Deep Learning**

### *7.1. Differences between the DNNs and ANNs*

Based on the number of neural networks, we distinguished between ANNs, which are neural networks consisting of a single hidden layer, and DNNs, which are neural

networks that have multiple hidden layers (as shown in Figure 9). This allows the network to understand and mimic more complex and abstract behaviours [38].



**Figure 9.** Graphical representation of an artificial neural network (ANN) and a deep neural network (DNN) [39].

### 7.2. Deep Neural Network (DNN) Classifiers

The DNN is a machine learning method with multiple hidden layers. It processes information from the input to the output, utilising weights and backpropagation to minimise the formation of errors. Adding more hidden layers improves the results obtained with DNNs, but also increases their computational and memory requirements as a consequence [40–43].

In the study published by Han et al., the ability of the ANNs and DNNs in evaluating drug formulations was assessed using 145 formulations with their reported dissolution times. Various factors, including the drug components, filler amount, and manufacturing parameters, were all taken into account. Both the ANNs and DNNs achieved a predictive value of 85.6% in the training set. The validation sets showed an 80% efficacy for ANNs and 85% for DNNs, respectively. However, on the crucial test set, DNNs significantly outperformed ANNs with an effectiveness of 85% compared to 80%, respectively. These results suggest that DNNs are better at predicting unknown data than ANNs [44].

Deep neural networks can be found in numerous healthcare sectors due to their effectiveness. They are currently applied from medical imaging, diagnosis, drug development, prognosis, and risk assessment, to remote monitoring and sports medicine [45–47]. The largest number of recent studies report the use of DNNs in the analysis of radiological images, among which include: models detecting apparent and non-apparent scaphoid fractures using only plain wrist radiographs, algorithms for COVID-19 detection from CXR images, models for mammography screening, and tools for the segmentation of intracerebral haemorrhage on CT scans [48–51]. DNNs are also involved in assessing endomyocardial biopsy data of patients with myocardial injury [52]. Others can identify hypertension on the basis of ballistocardiogram signals or aortic stenosis using audio files [53,54].

Models designed to make diagnoses based on the provided data are gaining popularity—these networks can learn patterns and extract valuable insights from large datasets, assisting in the identification of diseases, like cancer, diabetic retinopathy, and cardiovascular diseases [55–57].

### 7.3. Convolutional Neural Networks

Convolutional neural networks (CNNs) consist of input, spline, auxiliary, and output layers. They detect visual patterns obtained from raw image pixels using hidden layers.

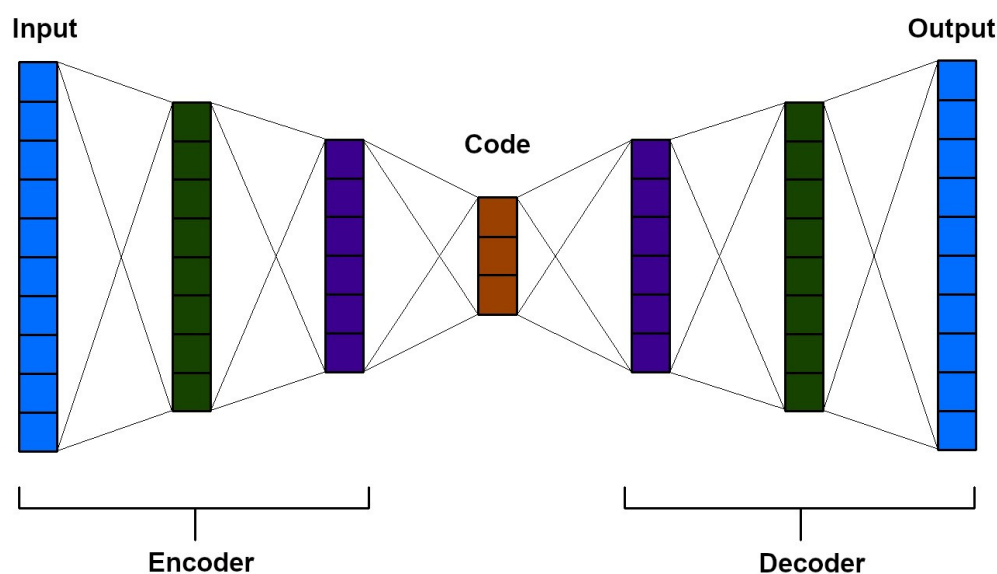
CNN models analyse receptive fields with filters [58]. Non-linear functions extract information about the image features. ‘Pooling’ reduces data and speeds up computation. This allows for the finding of similar features across the image for pattern analysis [59].

A study by Chamberlin et al. compared the performance of a CNN with expert radiologists in detecting pulmonary nodules and in determining the coronary artery calcium volume (CACV) on low-dose chest CT (LDCT) images. The study included 117 patients for pulmonary nodules and 96 for CACV, respectively. AI results demonstrated an excellent concordance with the radiologists (CACV ICC = 0.904, Cohen’s kappa for pulmonary nodules = 0.846) along with a high sensitivity/specificity (CACV: sensitivity = 0.929, specificity = 0.960; pulmonary nodules: sensitivity = 1, specificity = 0.708, respectively) [60].

Considering CNNs as a form of large-scale learning, it is therefore crucial to understand how to integrate the knowledge acquired from all datasets without putting too much effort into it. However, the performance of CNNs can be impacted through various factors, the most important of which include the selection of activation functions and the quantity of hidden layers. As a result, the accuracy of each CNN experiment fluctuates based on the sizes of the hidden layers that are chosen [61].

#### 7.4. Auto-Encoders (Unsupervised)

The auto-encoder (AE) neural network is an unsupervised learning model that reconstructs input images. It utilises an encoder to compress the data and a decoder to reconstruct the output layer (Figure 10). Auto-encoders are effective in data compression and visualisation, particularly in the initial learning phase of neural networks. They aid in identifying patterns in data series. Initial training with auto-encoders addresses potential issues, such as excessive parameters in the DNNs or disparities in gradient magnitudes between the higher and lower layers [62,63].



**Figure 10.** Graphical representation of the auto-encoder [64]. The light blue colour indicates the input and output layers. The dark green colour together with the dark blue colour indicates the internal layers of the encoder and decoder. Red, on the other hand, stands for code.

Auto-encoders are commonly applied in clustering, which involves grouping similar data together, particularly in large datasets. They can identify patterns in data arrangements, such as the sequence of natural numbers or the pattern of a geometric sequence.

Auto-encoders have been used for anomaly detection in brain MRIs. Baur et al. utilised pre-trained auto-encoders to model the normal white matter anatomy of the brain. By reconstructing images and comparing them to the standard model, anomalies and errors could be identified as outliers. This approach is beneficial for detecting uncommon

pathologies that may be missed from supervised learning. However, none of the employed autoencoder-based models were able to accurately restore the pathological equivalents of the baseline samples [65].

#### 7.5. Segmenting Neural Networks (e.g., UNET and Lung Segmentation)

Current medical image segmentation methods often rely on full CNNs with a U-shaped structure, such as UNET networks. UNET networks feature a symmetric encoder-decoder architecture with minimal connections. The encoder extracts deep features with large receptive fields through convolutional and down-sampling layers. The decoder then up-samples these features to match the input resolution, enabling pixel-level semantic prediction. Minimal connections are primarily used to combine high-resolution features and different scales at the end of the process, reducing the extent of data loss from down-sampling [66].

UNET networks have various applications, including face recognition in images, which are commonly used in smartphones. Shamim et al. proposed convUnet, a modified UNET architecture with additional convolutional layers in each decoder block to identify matte glass areas in lung computed tomography (CT) images. This enhancement significantly improved the segmentation performance for interconnected lung areas in CT imaging. This method shows promise for rapid COVID-19 diagnosis and quantification of infected lung regions [67].

#### 7.6. Generative Adversarial Networks

A generative adversarial network (GAN) consists of a generator and a discriminator. GANs learn to generate realistic data by competing against each other. They have been successfully applied in medicine for generating various types of medical images, such as mammograms, CT scans, and MRIs. This allows for the training of models that require diverse image data. GANs are also utilised in generating medical data, including electrocardiograms (ECGs), which can effectively train other models.

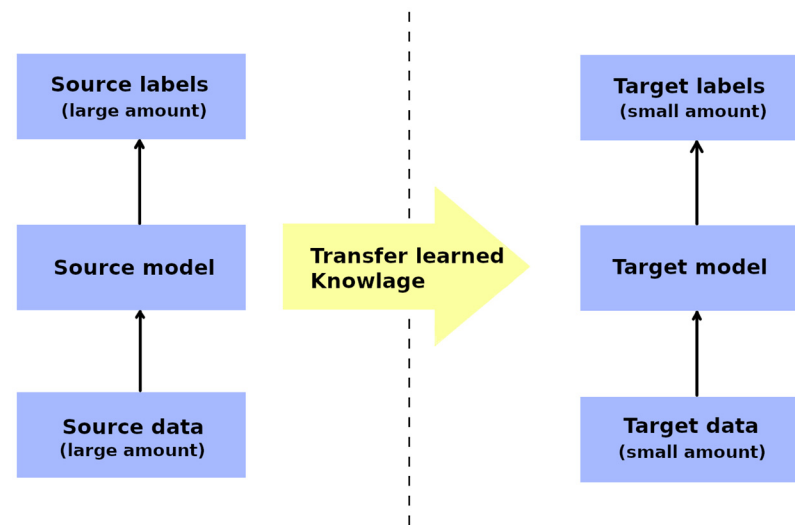
#### 7.7. Transfer Learning

Transfer learning is a machine learning technique that aims to enhance the performance of a target task by leveraging the knowledge or representations learned from a different but related source task (Figure 11). It involves the transfer of knowledge, skills, or features from a pre-trained model (source domain) to a new task (target domain) with limited labelled data [68]. However, transfer learning is not without its limitations and potential drawbacks. One notable limitation is the assumption that the source and target domains share some degree of similarity or relatedness. Furthermore, if the domains differ significantly, the transferred knowledge may not be relevant nor applicable to the target domain, resulting in a degraded performance. In such cases, the benefits of transfer learning diminish, and a domain-specific or task-specific model may be more appropriate [69].

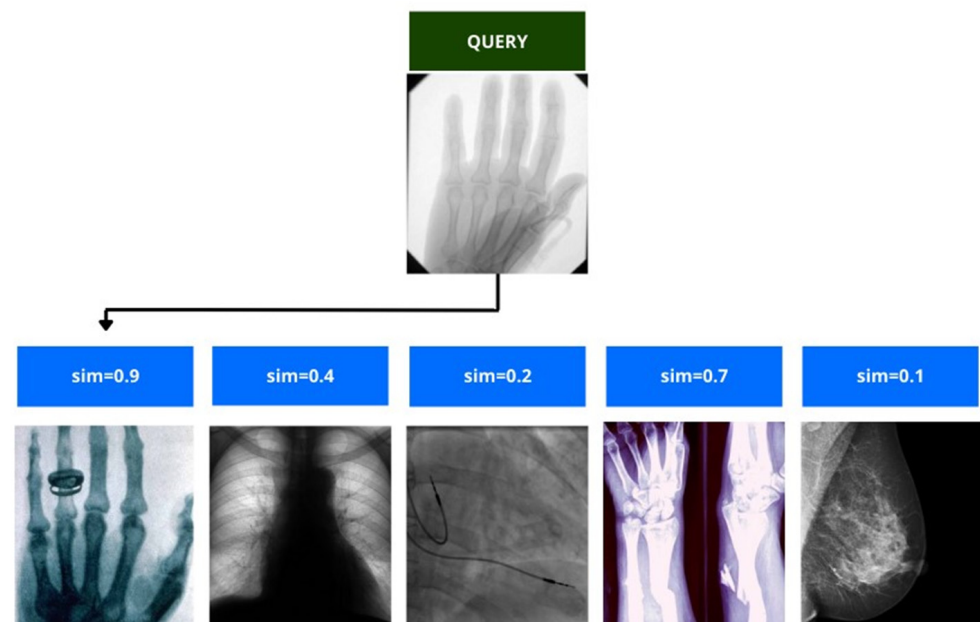
#### 7.8. Few-Shot Learning

Few-shot learning is a machine learning paradigm that deals with the task of learning new concepts or categories from only a few labelled examples. It addresses the challenge of training accurate models with limited labelled data, which is a common scenario in many real-world applications. Unlike traditional machine learning approaches that require large amounts of labelled data for each class, few-shot learning aims to generalise the knowledge obtained from a small support set to classify or generate instances from novel classes [70]. One significant challenge is the scarcity of labelled examples, which makes it difficult for models to capture the underlying patterns and variances in the data (Figure 12). This can lead to a limited generalization capability along with a poor performance on novel classes or instances which do not present in the support set. Another potential defect is the vulnerability of few-shot learning models to overfitting. With limited labelled data, these

models may be more prone to fitting noise or idiosyncrasies in the support set, resulting in a reduced performance when faced with new, unseen instances [71].



**Figure 11.** Graphical representation of the transfer learning technique.

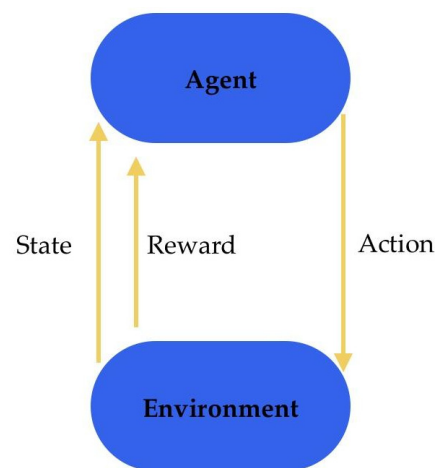


**Figure 12.** Simplified representation of the few-shot learning paradigm [72].

### 7.9. Deep Reinforcement Learning

Deep reinforcement learning (DRL) is a type of machine learning that enables machines to learn and make decisions under complex environments. It integrates the power of DNNs with the ability of reinforcement learning algorithms to learn through trial and error and to take actions that maximise a cumulative reward signal. The environment provides the agent with observations and rewards based on its actions (Figure 13). The agent is a learning entity that processes received observations through DNNs and selects actions based on the learned policy. DRL has been successfully applied in various domains, including in medicine. It also holds promise in other areas, such as medical diagnosis and decision making, clinical trial design, and robotic surgery.

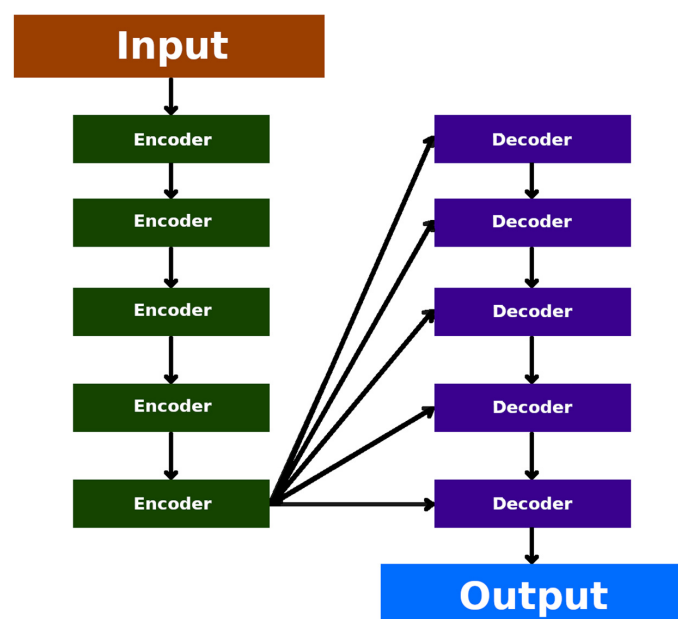




**Figure 13.** Graphical representation of deep reinforcement learning [73].

#### 7.10. Transformer Neural Networks

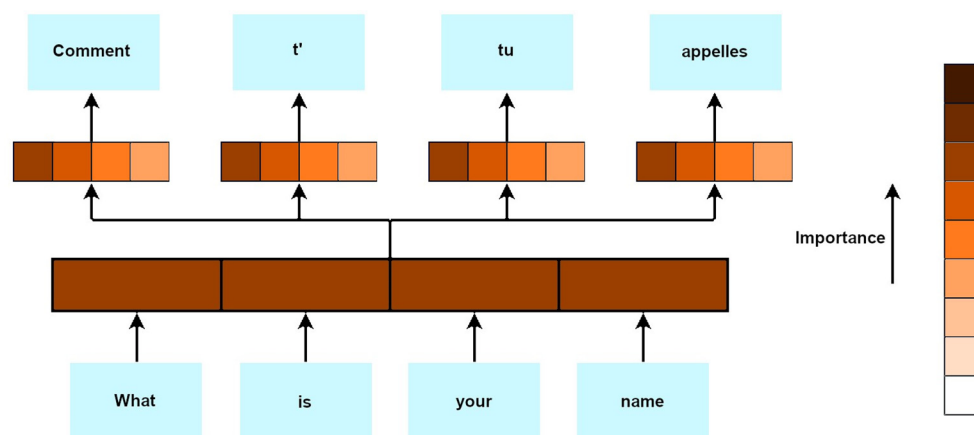
Transformer neural networks (TNN), also known as transformers, are powerful neural networks that have been widely used in natural language processing [36]. Transformers were developed to solve the problems of sequence-to-sequence transduction and neural machine translation. This includes speech recognition, text-to-speech transformation, etc. They consist of encoder–decoder layers and are trained through pre-training and fine-tuning. (Figure 14) Unlike the traditional recurrent neural networks (RNNs), which process inputs sequentially, transformers exploit a self-attention mechanism (also known as scaled dot-product attention) to weigh the importance of different elements in a sequence, enabling parallelisation and capturing long-range dependencies more effectively. In doing so, the model can thereby assign higher weights to relevant words or tokens and pay attention to them more effectively during processing. While transformers have been highly successful in natural language processing, they also have limitations. These include a high computational complexity, memory requirements for long sequences, a lack of explicit sequential order modelling, limited interpretability, data-intensive training, difficulty with contextual understanding, and challenges with out-of-vocabulary words.



**Figure 14.** Simplified diagram of TNN operation [74].

### 7.11. Attention Mechanism

The attention mechanism is a component of artificial intelligence (AI) models that enables them to focus on specific parts of input data that are deemed relevant for a given task. It allows the model to assign weights to the different elements of the input data, directing its attention to the most important information (Figure 15) [75].



**Figure 15.** Attention mechanism—simple diagram.

In medical imaging, attention mechanisms can be used to highlight specific regions or structures within an image. For instance, in a chest X-ray, attention can be directed to areas of potential pathology, thereby assisting radiologists in their diagnosis and improving the accuracy of automated image analysis algorithms.

Attention mechanisms can aid in clinical decision-making by directing the model's attention to the relevant information within patient records. When analyzing electronic health records (EHRs) or the medical literature, the model can focus on crucial clinical features, symptoms, or treatment options, thereby assisting healthcare professionals in making informed decisions [76].

The weakness of these attention mechanisms in AI is that they can be sensitive to input variations, resulting in the formation of different attention distributions for similar inputs. This sensitivity can lead to an instability and inconsistency in the model's attention patterns, thereby affecting its performance. Additionally, attention mechanisms may struggle to handle long-range dependencies in sequences, making it challenging to capture and incorporate information from the distant parts of the input sequence more effectively.

## 8. Examples of AI Applications in Medicine Approved by the US Food and Drug Administration (FDA)

- **Apple IRNF 2.0:** Apple Inc. developed the IRNF 2.0 software for the Apple Watch, utilising CNNs and machine learning algorithms. This software aims to identify cardiac rhythm disorders, particularly atrial fibrillation (AFib). The study conducted by Apple in 2021, with FDA approval, involved over 2500 participants and collected more than 3 million heart rate recordings. The algorithm successfully differentiated between AFib and non-AFib rhythms, with a sensitivity of 88.6% and a specificity of 99.3%. While effective, it is important to note that the app does not replace professional diagnosis nor target individuals that have already been diagnosed with AFib [77].
- **Ultromics:** An AI-powered solution detects heart failure with a 90% accuracy, specifically heart failure with preserved ejection fraction (HFpEF). It analyses LV images using an AI ML-based algorithm to accurately measure LV parameters, such as volumes, left ventricular ejection fraction (LVEF), and left ventricular longitudinal strain (LVLS). This software (EchoGo Core 2.0) also classifies echocardiographic views for quality control. AI readings are more consistent than manual readings, regardless of the image quality. Additionally, AI-derived LVEF and LVLS values have been significantly associated with mortality in-hospital and at final follow-up [78].

- **Aidoc:** Aidoc is an AI-based platform that permits a fast and accurate analysis of X-rays and CT scans. It detects conditions, like strokes, fractures, and cancerous lesions. Their advanced AI softwares (Aidoc software from 2.0 and above) help radiologists to prioritise their critical cases and expedite patient care. Aidoc has 12 FDA-approved tools, including softwares (Aidoc software from 2.0 and above) for analysing head CT images (detecting intracranial haemorrhage), chest CT studies (identifying aortic dissection), chest X-rays (flagging a suspected pneumothorax), and abdominal CT images (indicating suspected intra-abdominal free gas) [79].
- **Riverain Technologies:** This AI-based platform enables an accurate and efficient analysis of lung images for detecting and monitoring lung conditions, like cancer, COPD, and bronchial disease. It includes features such as CT Detect for measuring areas of interest, ClearRead CT Compare for comparing nodules across studies, and ClearRead CT Vessel Suppress for enhancing nodule visibility. This patented technology improves the accuracy and reading performance, and seamlessly integrates processed series with the original CT series for synchronised scrolling [80].

## 9. Discussion and Limitations

The development of artificial intelligence faces many challenges, and a lack of access to the relevant training data is a major problem in training AI models. Manual data labelling is costly, time consuming, and error prone [81]. It is important to ensure the reliability and interpretability of AI and data fusion methods, requiring regulation and ethical considerations. Research on the quality of healthcare services, privacy, and limitations in applying AI to healthcare is crucial [82]. With the development of artificial intelligence, the risk of hacking attacks on medical data, including patient data, is increasing. Data protection in the medical sector is becoming extremely significant, requiring advanced solutions and close cooperation between security specialists, medical IT specialists, and healthcare professionals. The use of AI to monitor and detect anomalies in IT networks and the continuous improvement of security features are key for minimizing the risk of hacking attacks and for protecting the patient data. An essential part of protecting patient data is the development of intelligent systems that use artificial intelligence to monitor and detect anomalies in IT networks. Artificial neural networks optimised for game theory can be used to detect anomalies and identify potential threats in real time [83].

Although machine learning is widely used, there are also some negative sides and limitations of its methods. Firstly, a dependency on large amounts of labelled data, meaning careful data preparation and evaluation, is a time-consuming process [61]. Insufficient data can lead to overfitting or poor generalization, and annotating such data is a challenging process that can lead to the inadequate performance of the model. Acquiring adequate data in medicine is even more challenging, as there are considerable privacy concerns, limited access to medical records, and the need for expert annotations during the learning process [84,85]. Ensuring data security remains a critical challenge, as data breaches, along with an unauthorised access or mishandling of patient data can present legal and ethical consequences, as well as lead to breaches in confidentiality and the erosion of patients trust. Deep neural networks, particularly with multiple hidden layers, which can be considered as black boxes, lack interpretability [86,87]. Understanding how the model arrives at its predictions or explaining the reasoning behind its decisions can be demanding. This lack of interpretability can limit their applicability in domains where transparency is crucial. Furthermore, the model's generalisation can be questionable, since DNNs can also learn biases present in the training data, which leads to biased predictions and decisions [88]. Additionally, applying DNNs to novel and unseen scenarios remains a significant challenge. The clinical adoption and validation of AI-based systems can be a complex process, as it requires integration with the existing infrastructure, acceptance by healthcare professionals, and validation against previously established standards [89]. On the other hand, there is a risk of overreliance on machine learning models, which can result in the potential neglect of clinical expertise and suboptimal patient care. Healthcare workers can be tempted to

heavily rely on model outputs without considering the context and patient characteristics. It is crucial to recognise and address these limitations to ensure the effective application of such systems in medicine.

## 10. Usability

Our study serves as a unique, comprehensive guide, akin to an AI encyclopedia, that explores and explains the concepts of machine learning, artificial neural networks (ANNs), and deep learning (DL). It was created to help other researchers and practitioners understand the concepts of these techniques and provide concrete examples of their applications, highlighting their potential to revolutionise healthcare by improving diagnostics, disease prediction, drug development, and personalised treatment approaches.

Our team of authors strongly believes that our work is a one-of-a-kind and relevant paper. We could not find other papers investigating our topic. Our review contains general information about AI, which was broken down into individual techniques that are used in it to familiarise readers with basic information on the subject. Other articles cover our selected topics in detail, often going into technical details or focusing on only one selected field of medicine. Only one of the papers we identified treated a similar topic, but its aim was different from ours. The review of Ahmad et al. presented a variety of approaches to the practical applications of AI in clinical practice that were extended with a theoretical elaboration that was broader than ours [90]. They provided a large number of examples, which were grouped according to the target specialty. We also provided numerous examples of different models' usage but lacked the division to certain specialties. In the aforementioned work, only applications of AI methods, in general, were presented—this research was not focused on how the model works and how to distinguish it from other models. However, in our work, the main focus, and a great advantage, was the incorporation of the most important basics along with the presentation of the information about the individual artificial intelligence algorithms with additional diagrams showing how they work supported by examples of their practical applications.

We firmly believe that our work is an excellent resource for beginners, serving an accessible starting point for those new to the field of AI in medicine. Serving as a compendium, it condenses a vast amount of information into a useful guide, making it an invaluable asset for researchers, practitioners, and all of those worldwide who are intrigued by the fusion of AI and medicine.

## 11. Summary and Future Research

Techniques using artificial intelligence present a promising tool employed in medical fields. Observing the speed of the development of these AI methods, it is certain that in the near future, the use of these methods will be an indispensable part of doctors' work, improving the diagnostic and therapeutic process. Research shows that AI methods are proving themselves not only as tools for identifying numerous diseases, regardless of the diagnosis carried out by specialists, but also as predictive tools for given disease entities, or as tools for observing changes in a patient's state of health. To date, some of these solutions have been approved by the FDA and have since been implemented for clinical use. Along with the developments above, doctors will be required to take a proactive approach to learning innovative medical technologies, which will undoubtedly make their work easier while improving medicine sectors.

The potential advancements of AI in medicine are vast and transformative. Future research should explore the integration of AI with emerging technologies, address ethical concerns, and focus on developing robust AI models that can enhance diagnosis, for example in X-rays, MRIs, and CT scans [91]. It can also consider factors, such as patient demographics, genetics, and comorbidities to optimise treatment strategies [92]. AI can optimise healthcare workflows by automating administrative tasks, predicting patient flow and hospital resource utilization, and providing real-time decision support for healthcare professionals [93]. Through pursuing these research directions, we can harness the full

potential of AI to improve patient outcomes, enhance healthcare efficiency, and drive innovations in the field of medicine. Future research in AI and medicine should also focus on developing interpretable and explainable AI models to enhance trust and transparency. This will enable healthcare professionals to understand and validate the reasoning behind AI-generated recommendations, thereby facilitating collaboration between humans and machines in clinical decision-making. Additionally, exploring the integration of AI with emerging technologies, such as blockchains, the Internet of Things (IoT), and augmented reality (AR), can open up new possibilities for secure and immersive healthcare experiences. Through continuously advancing AI research, we can unlock its full potential to revolutionise healthcare delivery and improve patient care.

**Author Contributions:** Conceptualization, J.K., S.K. and I.P.; methodology, J.K., M.J. and P.D.; data curation, J.K., W.B., A.L. and I.P.; writing—original draft preparation, J.K., K.B.-L., S.K., M.K., W.B., M.J., Ł.C., P.D., M.M., A.L. and I.P.; writing—review and editing, J.K., K.B.-L., S.K., M.K., W.B., M.J., Ł.C., P.D., M.M., A.L. and I.P.; visualization, J.K., K.B.-L. and Ł.C.; supervision, J.K., S.K. and K.B.-L.; project administration, J.K.; funding acquisition, J.K., Z.N., M.C. and K.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:** No new data were created or analysed in this study. Data sharing is not applicable to this article.

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

## References

1. Ward, T.M.; Mascagni, P.; Madani, A.; Padoy, N.; Perretta, S.; Hashimoto, D.A. Surgical Data Science and Artificial Intelligence for Surgical Education. *J. Surg. Oncol.* **2021**, *124*, 221–230. [CrossRef]
2. Mortani Barbosa, E.J.; Geftter, W.B.; Ghesu, F.C.; Liu, S.; Mailhe, B.; Mansoor, A.; Grbic, S.; Vogt, S. Automated Detection and Quantification of COVID-19 Airspace Disease on Chest Radiographs: A Novel Approach Achieving Expert Radiologist-Level Performance Using a Deep Convolutional Neural Network Trained on Digital Reconstructed Radiographs from Computed Tomography-Derived Ground Truth. *Investig. Radiol.* **2021**, *56*, 471–479. [CrossRef]
3. Gupta, M. Introduction to Data in Machine Learning. GeeksforGeeks. Available online: <https://www.geeksforgeeks.org/ml-introduction-data-machine-learning/> (accessed on 1 May 2023).
4. Dorfman, E. How Much Data Is Required for Machine Learning? Postindustria. Available online: <https://postindustria.com/how-much-data-is-required-for-machine-learning/> (accessed on 7 May 2023).
5. Patel, H. Data-Centric Approach vs. Model-Centric Approach in Machine Learning. MLOps Blog 2023. Available online: <https://neptune.ai/blog/data-centric-vs-model-centric-machine-learning> (accessed on 1 May 2023).
6. Brown, S. Machine Learning, Explained. MIT Sloan. Ideas Made to Matter. 2021. Available online: <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained> (accessed on 4 May 2023).
7. Christopher, A. K-Nearest Neighbor. Medium. The Startup. 2021. Available online: <https://medium.com/swlh/k-nearest-neighbor-ca2593d7a3c4> (accessed on 10 May 2023).
8. Hamed, A.; Sobhy, A.; Nassar, H. Accurate Classification of COVID-19 Based on Incomplete Heterogeneous Data Using a KNN Variant Algorithm. *Arab J. Sci. Eng.* **2021**, *46*, 8261–8272. [CrossRef]
9. Bellino, G.; Schiaffino, L.; Battisti, M.; Guerrero, J.; Rosado-Muñoz, A. Optimization of the KNN Supervised Classification Algorithm as a Support Tool for the Implantation of Deep Brain Stimulators in Patients with Parkinson's Disease. *Entropy* **2019**, *21*, 346. [CrossRef] [PubMed]
10. What Is Linear Regression? IBM. Available online: <https://www.ibm.com/topics/linear-regression> (accessed on 11 May 2023).
11. Garcia, J.M.V.; Bahloul, M.A.; Laleg-Kirati, T.-M. A Multiple Linear Regression Model for Carotid-to-Femoral Pulse Wave Velocity Estimation Based on Schrodinger Spectrum Characterization. In Proceedings of the 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Glasgow, UK, 11–15 July 2022; pp. 143–147.
12. Co to Jest Uczenie Maszynowe? Microsoft Azure. Available online: <https://azure.microsoft.com/pl-pl/resources/cloud-computing-dictionary/what-is-machine-learning-platform> (accessed on 11 May 2023).
13. Regresja Logistyczna. IBM. Available online: <https://www.ibm.com/docs/pl/spss-statistics/28.0.0?topic=regression-logistic> (accessed on 14 May 2023).



14. Kleinbaum, D.G.; Klein, M. Logistic Regression. In *Statistics for Biology and Health*; Springer: New York, NY, USA, 2010; ISBN 978-1-4419-1741-6.
15. Gruszczyński, M.; Witkowski, B.; Wiśniowski, A.; Szulc, A.; Owczarczuk, M.; Książek, M.; Bazyl, M. *Mikroekonometria. Modele i Metody Analizy Danych Indywidualnych*; Akademicka. Ekonomia; II; Wolters Kluwer Polska SA: Gdansk, Poland, 2012; ISBN 978-83-264-5184-3.
16. Naiwny Klasyfikator Bayesa. StatSoft Internetowy Podręcznik Statystyki. Available online: [https://www.statsoft.pl/textbook/stathome\\_stat.html?https%3A%2F%2Fwww.statsoft.pl%2Ftextbook%2Fgo\\_search.html%3Fq%3D%25bayesa](https://www.statsoft.pl/textbook/stathome_stat.html?https%3A%2F%2Fwww.statsoft.pl%2Ftextbook%2Fgo_search.html%3Fq%3D%25bayesa) (accessed on 2 May 2023).
17. Možina, M.; Demšar, J.; Kattan, M.; Zupan, B. Nomograms for Visualization of Naive Bayesian Classifier. In *Knowledge Discovery in Databases: PKDD 2004*; Lecture Notes in Computer Science; Boulicaut, J.-F., Esposito, F., Giannotti, F., Pedreschi, D., Eds.; Springer: Berlin/Heidelberg, Germany, 2004; Volume 3202, pp. 337–348. ISBN 978-3-540-23108-0.
18. Minsky, M. Steps toward Artificial Intelligence. *Proc. IRE* **1961**, *49*, 8–30. [[CrossRef](#)]
19. Zhou, S. Sparse SVM for Sufficient Data Reduction. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 5560–5571. [[CrossRef](#)] [[PubMed](#)]
20. Bordes, A.; Ertekin, S.; Weston, J.; Bottou, L. Fast Kernel Classifiers with Online and Active Learning. *J. Mach. Learn.* **2005**, *6*, 1579–1619.
21. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
22. Noble, W.S. What Is a Support Vector Machine? *Nat. Biotechnol.* **2006**, *24*, 1565–1567. [[CrossRef](#)]
23. Winters-Hilt, S.; Merat, S. SVM Clustering. *BMC Bioinform.* **2007**, *8*, S18. [[CrossRef](#)]
24. Huang, S.; Cai, N.; Pacheco, P.P.; Narrandes, S.; Wang, Y.; Xu, W. Applications of Support Vector Machine (SVM) Learning in Cancer Genomics. *Cancer Genom. Proteom.* **2018**, *15*, 41–51. [[CrossRef](#)]
25. Zhang, J.; Xu, J.; Hu, X.; Chen, Q.; Tu, L.; Huang, J.; Cui, J. Diagnostic Method of Diabetes Based on Support Vector Machine and Tongue Images. *BioMed Res. Int.* **2017**, *2017*, 7961494. [[CrossRef](#)] [[PubMed](#)]
26. Schapire, R.E.; Freund, Y. *Boosting: Foundations and Algorithms*; Adaptive Computation and Machine Learning Series; MIT Press: Cambridge, MA, USA, 2012; ISBN 978-0-262-01718-3.
27. Li, S.; Zeng, Y.; Chapman, W.C.; Erfanzadeh, M.; Nandy, S.; Mutch, M.; Zhu, Q. Adaptive Boosting (AdaBoost)-based Multiwave-length Spatial Frequency Domain Imaging and Characterization for Ex Vivo Human Colorectal Tissue Assessment. *J. Biophotonics* **2020**, *13*, e201960241. [[CrossRef](#)]
28. Hatwell, J.; Gaber, M.M.; Atif Azad, R.M. Ada-WHIPS: Explaining AdaBoost Classification with Applications in the Health Sciences. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 250. [[CrossRef](#)] [[PubMed](#)]
29. Baniasadi, A.; Rezaeairad, S.; Zare, H.; Ghassemi, M.M. Two-Step Imputation and AdaBoost-Based Classification for Early Prediction of Sepsis on Imbalanced Clinical Data. *Crit. Care Med.* **2021**, *49*, e91–e97. [[CrossRef](#)] [[PubMed](#)]
30. Takemura, A.; Shimizu, A.; Hamamoto, K. Discrimination of Breast Tumors in Ultrasonic Images Using an Ensemble Classifier Based on the AdaBoost Algorithm with Feature Selection. *IEEE Trans. Med. Imaging* **2010**, *29*, 598–609. [[CrossRef](#)]
31. Salcedo-Sanz, S.; Pérez-Aracil, J.; Ascenso, G.; Del Ser, J.; Casillas-Pérez, D.; Kadow, C.; Fister, D.; Barriopedro, D.; García-Herrera, R.; Restelli, M.; et al. Analysis, Characterization, Prediction and Attribution of Extreme Atmospheric Events with Machine Learning: A Review. *arXiv* **2022**, arXiv:2207.07580. [[CrossRef](#)]
32. Moore, A.; Bell, M. XGBoost, A Novel Explainable AI Technique, in the Prediction of Myocardial Infarction: A UK Biobank Cohort Study. *Clin. Med. Insights Cardiol.* **2022**, *16*, 117954682211336. [[CrossRef](#)]
33. Wang, X.; Zhu, T.; Xia, M.; Liu, Y.; Wang, Y.; Wang, X.; Zhuang, L.; Zhong, D.; Zhu, J.; He, H.; et al. Predicting the Prognosis of Patients in the Coronary Care Unit: A Novel Multi-Category Machine Learning Model Using XGBoost. *Front. Cardiovasc. Med.* **2022**, *9*, 764629. [[CrossRef](#)]
34. Subha Ramakrishnan, M.; Ganapathy, N. Extreme Gradient Boosting Based Improved Classification of Blood-Brain-Barrier Drugs. In *Studies in Health Technology and Informatics*; Séroussi, B., Weber, P., Dhombres, F., Grouin, C., Liebe, J.-D., Pelayo, S., Pinna, A., Rance, B., Sacchi, L., Ugon, A., et al., Eds.; IOS Press: Amsterdam, The Netherlands, 2022; ISBN 978-1-64368-284-6.
35. Inoue, T.; Ichikawa, D.; Ueno, T.; Cheong, M.; Inoue, T.; Whetstone, W.D.; Endo, T.; Nizuma, K.; Tominaga, T. XGBoost, a Machine Learning Method, Predicts Neurological Recovery in Patients with Cervical Spinal Cord Injury. *Neurotrauma Rep.* **2020**, *1*, 8–16. [[CrossRef](#)]
36. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. In *Proceedings of the NIPS 2017, Long Beach, CA, USA, 4–9 December 2017*. [[CrossRef](#)]
37. Shao, R.; Shi, Z.; Yi, J.; Chen, P.-Y.; Hsieh, C.-J. On the Adversarial Robustness of Vision Transformers. *arXiv* **2021**, arXiv:2103.15670. [[CrossRef](#)]
38. Qureshi, J. What Is the Difference between Neural Networks and Deep Neural Networks? Quora 2018. Available online: <https://www.quora.com/What-is-the-difference-between-neural-networks-and-deep-neural-networks> (accessed on 3 May 2023).
39. Jeffrey, C. Explainer: What Is Machine Learning? TechSpot 2020. Available online: <https://www.techspot.com/article/2048-machine-learning-explained/> (accessed on 3 May 2023).
40. McBee, M.P.; Awan, O.A.; Colucci, A.T.; Ghobadi, C.W.; Kadom, N.; Kansagra, A.P.; Tridandapani, S.; Auffermann, W.F. Deep Learning in Radiology. *Acad. Radiol.* **2018**, *25*, 1472–1480. [[CrossRef](#)] [[PubMed](#)]



41. Chan, H.-P.; Samala, R.K.; Hadjiiski, L.M.; Zhou, C. Deep Learning in Medical Image Analysis. In *Deep Learning in Medical Image Analysis*; Advances in Experimental Medicine and Biology; Lee, G., Fujita, H., Eds.; Springer International Publishing: Cham, Switzerland, 2020; Volume 1213, pp. 3–21. ISBN 978-3-030-33127-6.
42. Kriegeskorte, N.; Golan, T. Neural Network Models and Deep Learning. *Curr. Biol.* **2019**, *29*, R231–R236. [CrossRef] [PubMed]
43. Bajić, F.; Orel, O.; Habijan, M. A Multi-Purpose Shallow Convolutional Neural Network for Chart Images. *Sensors* **2022**, *22*, 7695. [CrossRef]
44. Han, R.; Yang, Y.; Li, X.; Ouyang, D. Predicting Oral Disintegrating Tablet Formulations by Neural Network Techniques. *Asian J. Pharm. Sci.* **2018**, *13*, 336–342. [CrossRef]
45. Egger, J.; Gsaxner, C.; Pepe, A.; Pomykala, K.L.; Jonske, F.; Kurz, M.; Li, J.; Kleesiek, J. Medical Deep Learning—A Systematic Meta-Review. *Comput. Methods Programs Biomed.* **2022**, *221*, 106874. [CrossRef]
46. Jafari, R.; Spincemaille, P.; Zhang, J.; Nguyen, T.D.; Luo, X.; Cho, J.; Margolis, D.; Prince, M.R.; Wang, Y. Deep Neural Network for Water/Fat Separation: Supervised Training, Unsupervised Training, and No Training. *Magn. Reson. Med.* **2021**, *85*, 2263–2277. [CrossRef]
47. Hou, J.-J.; Tian, H.-L.; Lu, B. A Deep Neural Network-Based Model for Quantitative Evaluation of the Effects of Swimming Training. *Comput. Intell. Neurosci.* **2022**, *2022*, 5508365. [CrossRef]
48. Singh, A.; Ardakani, A.A.; Loh, H.W.; Anamika, P.V.; Acharya, U.R.; Kamath, S.; Bhat, A.K. Automated Detection of Scaphoid Fractures Using Deep Neural Networks in Radiographs. *Eng. Appl. Artif. Intell.* **2023**, *122*, 106165. [CrossRef]
49. Gülmez, B. A Novel Deep Neural Network Model Based Xception and Genetic Algorithm for Detection of COVID-19 from X-Ray Images. *Ann. Oper. Res.* **2022**. [CrossRef]
50. Tsai, K.-J.; Chou, M.-C.; Li, H.-M.; Liu, S.-T.; Hsu, J.-H.; Yeh, W.-C.; Hung, C.-M.; Yeh, C.-Y.; Hwang, S.-H. A High-Performance Deep Neural Network Model for BI-RADS Classification of Screening Mammography. *Sensors* **2022**, *22*, 1160. [CrossRef]
51. Sharrock, M.F.; Mould, W.A.; Ali, H.; Hildreth, M.; Awad, I.A.; Hanley, D.F.; Muschelli, J. 3D Deep Neural Network Segmentation of Intracerebral Hemorrhage: Development and Validation for Clinical Trials. *Neuroinform* **2021**, *19*, 403–415. [CrossRef] [PubMed]
52. Jiao, Y.; Yuan, J.; Sodimu, O.M.; Qiang, Y.; Ding, Y. Deep Neural Network-Aided Histopathological Analysis of Myocardial Injury. *Front. Cardiovasc. Med.* **2022**, *8*, 724183. [CrossRef]
53. Rajput, J.S.; Sharma, M.; Kumar, T.S.; Acharya, U.R. Automated Detection of Hypertension Using Continuous Wavelet Transform and a Deep Neural Network with Ballistocardiography Signals. *IJERPH* **2022**, *19*, 4014. [CrossRef] [PubMed]
54. Voigt, I.; Boeckmann, M.; Bruder, O.; Wolf, A.; Schmitz, T.; Wieneke, H. A Deep Neural Network Using Audio Files for Detection of Aortic Stenosis. *Clin. Cardiol.* **2022**, *45*, 657–663. [CrossRef]
55. Ma, L.; Yang, T. Construction and Evaluation of Intelligent Medical Diagnosis Model Based on Integrated Deep Neural Network. *Comput. Intell. Neurosci.* **2021**, *2021*, 7171816. [CrossRef]
56. Ragab, M.; AL-Ghamdi, A.S.A.-M.; Fakieh, B.; Choudhry, H.; Mansour, R.F.; Koundal, D. Prediction of Diabetes through Retinal Images Using Deep Neural Network. *Comput. Intell. Neurosci.* **2022**, *2022*, 7887908. [CrossRef]
57. Min, J.K.; Yang, H.-J.; Kwak, M.S.; Cho, C.W.; Kim, S.; Ahn, K.-S.; Park, S.-K.; Cha, J.M.; Park, D.I. Deep Neural Network-Based Prediction of the Risk of Advanced Colorectal Neoplasia. *Gut Liver* **2021**, *15*, 85–91. [CrossRef] [PubMed]
58. Anwar, S.M.; Majid, M.; Qayyum, A.; Awais, M.; Alnowami, M.; Khan, M.K. Medical Image Analysis Using Convolutional Neural Networks: A Review. *J. Med. Syst.* **2018**, *42*, 226. [CrossRef]
59. Mohamed, E.A.; Gaber, T.; Karam, O.; Rashed, E.A. A Novel CNN Pooling Layer for Breast Cancer Segmentation and Classification from Thermograms. *PLoS ONE* **2022**, *17*, e0276523. [CrossRef]
60. Chamberlin, J.; Kocher, M.R.; Waltz, J.; Snoddy, M.; Stringer, N.F.C.; Stephenson, J.; Sahbaee, P.; Sharma, P.; Rapaka, S.; Schoepf, U.J.; et al. Automated Detection of Lung Nodules and Coronary Artery Calcium Using Artificial Intelligence on Low-Dose CT Scans for Lung Cancer Screening: Accuracy and Prognostic Value. *BMC Med.* **2021**, *19*, 55. [CrossRef] [PubMed]
61. Alajanbi, M.; Malerba, D.; Liu, H. Distributed Reduced Convolution Neural Networks. *Mesopotamian J. Big Data* **2021**, *2021*, 26–29. [CrossRef]
62. Le, Q.V. A Tutorial on Deep Learning Part 2: Autoencoders, Convolutional Neural Networks and Recurrent Neural Networks. *Google Brain* **2015**, *20*, 1–20.
63. Ren, L.; Sun, Y.; Cui, J.; Zhang, L. Bearing Remaining Useful Life Prediction Based on Deep Autoencoder and Deep Neural Networks. *J. Manuf. Syst.* **2018**, *48*, 71–77. [CrossRef]
64. Dertat, A. Applied Deep Learning-Part 3: Autoencoders. Medium. Towards Data Science. 2017. Available online: <https://towardsdatascience.com/applied-deep-learning-part-3-autoencoders-1c083af4d798> (accessed on 3 May 2023).
65. Baur, C.; Denner, S.; Wiestler, B.; Navab, N.; Albarqouni, S. Autoencoders for Unsupervised Anomaly Segmentation in Brain MR Images: A Comparative Study. *Med. Image Anal.* **2021**, *69*, 101952. [CrossRef]
66. Cao, H.; Wang, Y.; Chen, J.; Jiang, D.; Zhang, X.; Tian, Q.; Wang, M. Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation. *arXiv* **2022**, arXiv:2105.05537. [CrossRef]
67. Shamim, S.; Awan, M.J.; Mohd Zain, A.; Naseem, U.; Mohammed, M.A.; Garcia-Zapirain, B. Automatic COVID-19 Lung Infection Segmentation through Modified Unet Model. *J. Healthc. Eng.* **2022**, *2022*, 6566982. [CrossRef]
68. Zhang, A.; Wang, H.; Li, S.; Cui, Y.; Liu, Z.; Yang, G.; Hu, J. Transfer Learning with Deep Recurrent Neural Networks for Remaining Useful Life Estimation. *Appl. Sci.* **2018**, *8*, 2416. [CrossRef]

69. Rios, A.; Kavuluru, R. Neural Transfer Learning for Assigning Diagnosis Codes to EMRs. *Artif. Intell. Med.* **2019**, *96*, 116–122. [CrossRef]
70. Snell, Swersky, K.; Zemel, R.S. Prototypical Networks for Few-Shot Learning. In Proceedings of the NIPS 2017, Long Beach, CA, USA, 4–9 December 2017. [CrossRef]
71. Wang, Y.; Wu, X.-M.; Li, Q.; Gu, J.; Xiang, W.; Zhang, L.; Li, V.O.K. Large Margin Few-Shot Learning. *arXiv* **2018**, arXiv:1807.02872.
72. Berger, M.; Yang, Q.; Maier, A. X-ray Imaging. In *Medical Imaging Systems; Lecture Notes in Computer Science*; Maier, A., Steidl, S., Christlein, V., Hornegger, J., Eds.; Springer International Publishing: Cham, Switzerland, 2018; Volume 11111, pp. 119–145. ISBN 978-3-319-96519-2.
73. Nie, M.; Chen, D.; Wang, D. Reinforcement Learning on Graphs: A Survey. *arXiv* **2022**, arXiv:2204.06127. [CrossRef]
74. Giacaglia, G. How Transformers Work. The Neural Network Used by Open AI and DeepMind. Towards Data Science 2019. Available online: <https://towardsdatascience.com/transformers-141e32e69591> (accessed on 1 May 2023).
75. Luo, X.; Gandhi, P.; Zhang, Z.; Shao, W.; Han, Z.; Chandrasekaran, V.; Turzhitsky, V.; Bali, V.; Roberts, A.R.; Metzger, M.; et al. Applying Interpretable Deep Learning Models to Identify Chronic Cough Patients Using EHR Data. *Comput. Methods Programs Biomed.* **2021**, *210*, 106395. [CrossRef] [PubMed]
76. Li, L.; Zhao, J.; Hou, L.; Zhai, Y.; Shi, J.; Cui, F. An Attention-Based Deep Learning Model for Clinical Named Entity Recognition of Chinese Electronic Medical Records. *BMC Med. Inform. Decis. Mak.* **2019**, *19*, 235. [CrossRef] [PubMed]
77. Available online. Available online: [https://www.accessdata.fda.gov/cdrh\\_docs/pdf21/K212516.pdf](https://www.accessdata.fda.gov/cdrh_docs/pdf21/K212516.pdf) (accessed on 5 May 2023).
78. Asch, F.M.; Descamps, T.; Sarwar, R.; Karagodin, I.; Singulane, C.C.; Xie, M.; Tucay, E.S.; Tude Rodrigues, A.C.; Vasquez-Ortiz, Z.Y.; Monaghan, M.J.; et al. Human versus Artificial Intelligence–Based Echocardiographic Analysis as a Predictor of Outcomes: An Analysis from the World Alliance Societies of Echocardiography COVID Study. *J. Am. Soc. Echocardiogr.* **2022**, *35*, 1226–1237.e7. [CrossRef]
79. Aidoc. Available online: <https://www.aidoc.com/solutions/radiology/> (accessed on 5 May 2023).
80. Riverain Technologies. Available online: <https://www.riveraintech.com/clearread-ai-solutions/clearread-ct/> (accessed on 5 May 2023).
81. Alzubaidi, L.; Bai, J.; Al-Sabaawi, A.; Santamaría, J.; Albahri, A.S.; Al-dabbagh, B.S.N.; Fadhel, M.A.; Manoufali, M.; Zhang, J.; Al-Timemy, A.H.; et al. A Survey on Deep Learning Tools Dealing with Data Scarcity: Definitions, Challenges, Solutions, Tips, and Applications. *J. Big Data* **2023**, *10*, 46. [CrossRef]
82. Albahri, A.S.; Duham, A.M.; Fadhel, M.A.; Alnoor, A.; Baqer, N.S.; Alzubaidi, L.; Albahri, O.S.; Alamoodi, A.H.; Bai, J.; Salhi, A.; et al. A Systematic Review of Trustworthy and Explainable Artificial Intelligence in Healthcare: Assessment of Quality, Bias Risk, and Data Fusion. *Inf. Fusion* **2023**, *96*, 156–191. [CrossRef]
83. Hephzipah, J.J.; Vallem, R.R.; Sheela, M.S.; Dhanalakshmi, G. An Efficient Cyber Security System Based on Flow-Based Anomaly Detection Using Artificial Neural Network. *Mesopotamian J. Cybersecur.* **2023**, *2023*, 48–56. [CrossRef]
84. Oliver, M.; Renou, A.; Allou, N.; Moscatelli, L.; Ferdynus, C.; Allyn, J. Image Augmentation and Automated Measurement of Endotracheal-Tube-to-Carina Distance on Chest Radiographs in Intensive Care Unit Using a Deep Learning Model with External Validation. *Crit. Care* **2023**, *27*, 40. [CrossRef]
85. Moon, J.-H.; Hwang, H.-W.; Yu, Y.; Kim, M.-G.; Donatelli, R.E.; Lee, S.-J. How Much Deep Learning Is Enough for Automatic Identification to Be Reliable? *Angle Orthod.* **2020**, *90*, 823–830. [CrossRef]
86. Albrecht, T.; Slabaugh, G.; Alonso, E.; Al-Arif, S.M.R. Deep Learning for Single-Molecule Science. *Nanotechnology* **2017**, *28*, 423001. [CrossRef]
87. Yang, Z.; Zhang, A.; Sudjianto, A. Enhancing Explainability of Neural Networks Through Architecture Constraints. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 2610–2621. [CrossRef] [PubMed]
88. Stock, P.; Cisse, M. ConvNets and ImageNet Beyond Accuracy: Understanding Mistakes and Uncovering Biases. *arXiv* **2017**, arXiv:1711.11443. [CrossRef]
89. Zheng, Q.; Yang, L.; Zeng, B.; Li, J.; Guo, K.; Liang, Y.; Liao, G. Artificial Intelligence Performance in Detecting Tumor Metastasis from Medical Radiology Imaging: A Systematic Review and Meta-Analysis. *EClinicalMedicine* **2021**, *31*, 100669. [CrossRef] [PubMed]
90. Ahmad, Z.; Rahim, S.; Zubair, M.; Abdul-Ghafar, J. Artificial Intelligence (AI) in Medicine, Current Applications and Future Role with Special Emphasis on Its Potential and Promise in Pathology: Present and Future Impact, Obstacles Including Costs and Acceptance among Pathologists, Practical and Philosophical Considerations. A Comprehensive Review. *Diagn. Pathol.* **2021**, *16*, 24. [CrossRef]
91. Syed, A.; Zoga, A. Artificial Intelligence in Radiology: Current Technology and Future Directions. *Semin. Musculoskelet. Radiol.* **2018**, *22*, 540–545. [CrossRef]
92. McDougall, R.J. Computer Knows Best? The Need for Value-Flexibility in Medical AI. *J. Med. Ethics* **2019**, *45*, 156–160. [CrossRef]
93. Dave, M.; Patel, N. Artificial Intelligence in Healthcare and Education. *Br. Dent. J.* **2023**, *234*, 761–764. [CrossRef]

**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.