Introduction to Machine Learning in Medicine

Rossana Buongiorno, Claudia Caudai, Sara Colantonio and Danila Germanese

Abstract Introduction of Machine Learning is certainly a revolutionary outbreak, both in the medical field and in many other research fields. The opportunities ML opens up are many, it can be used to analyze medical images, to relate clinical data, to identify lesions, to support diagnosis, to make predictions on prognosis and disease evolution, but it must be used carefully and competently so as not to run into wrong applications and consequently in misleading results. Machine Learning is represented by a series of algorithms which, through a learning process, are able to develop strategies to analyse, classify and evaluate many types of data. However, there are important points on which to pay close attention, such as the preparation of the learning dataset, the choice of the most suitable algorithm for the aimed task, the evaluation of the results, the weighing of the computing resources available. There are also some aspects related to Artificial Intelligence that need to be carefully considered, such as evaluation of the developed model, explainability, interpretability, reproducibility of the outputs, and all the legal and ethical issues that the nature of these approaches generates. Taking the necessary precautions for use, however, it cannot be denied that the application of ML in the medical field opens up very interesting and revolutionary perspectives, especially in the field of precision and personalized medicine.

1 Introduction

The increasing availability of patient-related data is driving new research trends addressing new personalised prediction and disease management. Nevertheless, the complexity of such an analysis makes it necessary the use of cognitive augmentation in the form of Artificial Intelligence (AI) systems [1, 2].

Rossana Buongiorno, Claudia Caudai, Sara Colantonio and Danila Germanese

Institute of Information Science and Technologies "A. Faedo" (ISTI), Pisa, Italy e-mail: rossana.buongiorno@isti.cnr.it

modern AI techniques have considerable potential to exploit complex medical data towards an improvement of current healthcare.

Many of the advances in this field are tied to progress in a subdomain of AI research known as *Machine Learning* (*ML*).

The scientist Arthur Lee Samuel was the first to introduce the term Machine Learning in 1959. He created the *checkers player*, a program designed to develop its own logic while playing and self-improve its performance.

ML algorithms are fruitfully applied in many research fields involving the most varied applications [3, 4, 5] . As far as applications in the medical and clinical fields are concerned, many uses, advantages and opportunities can be enumerated [6, 7, 8].

Nevertheless, ML is not yet used in the medical field as a direct diagnostic tool, but exclusively as a support to the diagnosis, because human supervision is still indispensable and probably will remain so for a long time. A vast trend is that of the analysis and interpretation of medical images, such as, for example, X-ray, US, echo, RM, using ML algorithms. Another widespread use is the prediction of the prognosis associated with particular diseases, such as different types of cancer. ML algorithms can also help automatically select patients suitable for particular clinical pathways and experimental trials. The analysis of the ML anatomical features can also be very useful in the real-time and non-real-time support of surgery, with the aim of promptly identifying unexpected critical or irregular anatomical structures. Other uses also include drug discovery, genomic screening, cardio-vascular tasks and epidemiology. In Table 1 many examples of medical applications are summarized.

In this chapter, we aim to describe, as simply as possible, what Machine Learning is and how it is possible to use it fruitfully in the medical field. In section 2, we describe the flow of a learning algorithm, in section 3, instead we report the main ML techniques and their most widespread clinical applications. In section 4, however, we briefly address some highly interesting issues (i.e. model evaluation, explanability, reproducibility, sharing, ethical and legal problems) and the great challenges of precision medicine and personalised medicine, which are increasingly handy thanks to Machine Learning.

2 What is Machine Learning?

Machine learning is a branch of AI which comprises elements of mathematics, statistics, and computer science. The term machine learning describes the ability of an algorithm to "learn" from data by recognising patterns and making inferences and/or predictions of future events with minimal human intervention and without explicit programming. In other words, the results produced by ML algorithms are inferences made from complex statistical analyses of adequately large datasets, expressed as the likelihood of a relationship between variables [9]. Furthermore, machine learning methods improve their performance adaptively as the number of examples from which they learn increase.

Table 1 Main medical applications for Machine Learning methods.

Applications	Description	Ref.
Cardio Vascular Tasks	Cardiovascular risk prediction.	2018 [59]
	Hyper-Myocardial infarction prediction.	2016 [60]
	Coronary artery disease detection.	2018 [61]
	Heart failure prediction.	2011 [36]
	Stroke risk prediction.	
	Cardiac arrhythmias evaluation.	
Human Cancer Imaging	Cancer detection on CT-scan, X-ray, US, MRI images.	2018 [62]
	Lesion segmentation.	2016 [57]
	Regression on cancer grade.	2019 [20]
	Cancer risk evaluation.	2022 [63]
	Cancer evolution prediction.	2021 [64]
	Cancer classification.	2019 [65]
Cancer Genomics	Cancer genomics classification.	2007 [28]
	Identification of pathogenic variants.	2013 [26]
	Detection of oncogenic states.	2014 [66]
	Bioactivity prediction.	2003 [33]
	Cancer cell-line specific interactions.	2014 [67]
	Clinical trials monitoring.	2021 [68]
Functional Genomics	Prediction of protein secondary structure.	2001 [69]
	Prediction of epigenomic marks.	2011 [48]
	Transcriptional and post-transcriptional regulation	2018 [70]
	investigation.	2015 [71]
	Transcription biological rules detection.	2020 [72]
	Protein Mass Spectrometry classification.	2021 [73]
Metabolic Disorders	Metabolic syndrome risk prediction	2021 [22]
	Alignment of metabolomic peaks.	2014 [74]
	Metabolites classification.	2017 [75]
	Kinetic metabolic modelling.	2016 [76]
	Prediction of estrogen receptor status.	2022 [77]
Prognostic Prediction	Survival Prediction.	2018 [78]
	Disease prognosis.	2019 [79]
	Trials outcome prediction.	2019 [80]
	Identification of risk factors.	2016 [81]
	Disease recurrence prediction.	2018 [82]
Drug Discovery	Drug target identification.	2018 [83]
	Target druggability prediction.	2016 [84]
	Splice variants classification.	2015 [85]
	Anticancer drugs prioritization.	2019 [86]

Over the past 5 years, machine-learned tools have demonstrated visible successes in the medical field, and in particular in disease detection, patient monitoring,

prognosis prediction, surgical assistance, patient care, and systems management, by supporting complex clinical decision-making [10, 11].

A wide variety of machine learning algorithms are in use today. The choice of a particular model for a given problem is determined by the characteristics of the data as well as the type of desired outcome. The majority of ML algorithms can be categorised into three types learning techniques: supervised learning, unsupervised learning, reinforcement learning (Fig. 1).
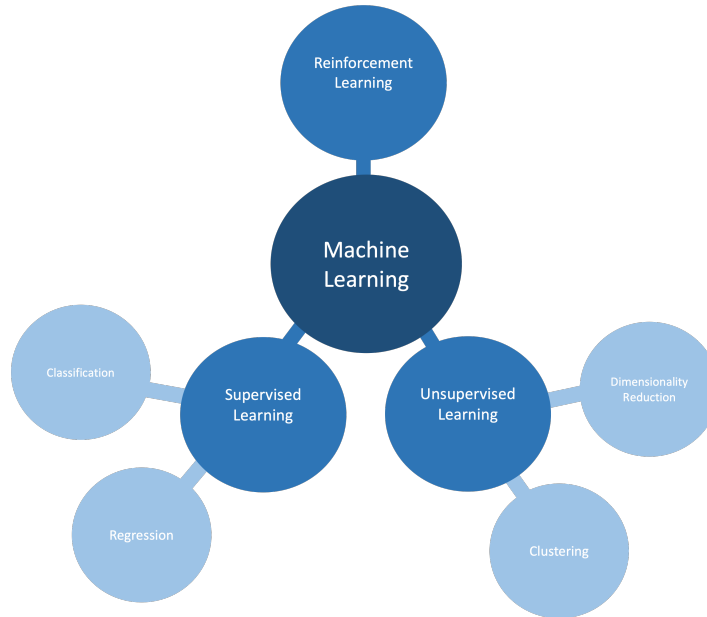

Fig. 1 ML algorithms are generally categorised as supervised learning, unsupervised learning, reinforcement learning

However, any type of algorithm consists of a series of key steps, as shown in Fig.2:

- Data collection: as machines learn from the data that one gives them, it is of the utmost importance to collect data from reliable sources. The higher the quality of the data, the higher the accuracy of the developed model. Good data are relevant, contain very few missing and repeated values, and have a good representation of the various subcategories/classes present. Noisy or incorrect data will clearly reduce the effectiveness of the model.
- Data pre-processing: after collecting data, they have to be correctly prepared. For instance, they may be randomised to make sure that data are evenly distributed, and the ordering does not affect the learning process. Also, data may be cleaned by removing missing/duplicate values, by converting data type etc. Finally, the cleaned data should be split into two sets: a training set and a testing set. The

training set is the set the model learns from. The testing set is used to check the accuracy of the trained model.

- Model training: the prepared data are analysed, elaborated and interpreted by the machine learning model to find patterns and make predictions. Over time, with training, the model gets better at learning from the data and inferring.
- Model testing: after training the model, its performances have to be checked. This is done by testing the accuracy and the speed of the model on previously unseen data (Testing Set).
- Model improving: this step is also known as Parameter Tuning. Once created and evaluated the model, the tuning of model parameters allows for improving the accuracy of the model itself. Parameters are the variables in the model that better fit the relationship between the data. At certain values of the parameters set, the accuracy may reach the maximum. Parameter tuning refers to finding these values.
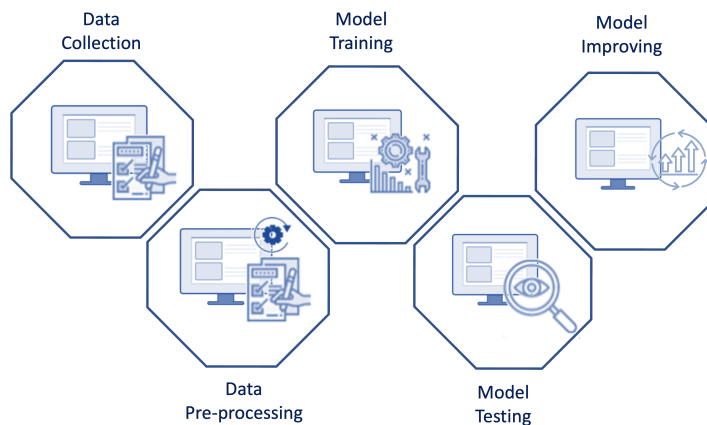
Fig. 2 Core steps of any type of ML approach

In the following sections the three main types of learning are described: supervised, unsupervised and reinforcement learning.

3 Principal ML Algorithms

Machine learning concerns three main types of algorithms: supervised learning, unsupervised learning and reinforcement learning. The difference between them is defined by how each algorithm learns the data to make predictions.

Supervised Machine Learning

Supervised learning refers to approaches in which a model is trained on a set of numerical inputs (or features, or predictors) which are associated with known outcomes (also referred to as ground truth, or prior knowledge).

As reported in Fig. 3, the goal in the first stage of learning is to best approximate the relationship between input and output observable in the data. In the validation step, the model is iteratively improved to reduce the error of prediction using optimisation techniques: in other words, the learning algorithm iteratively compares its predictions with the correct output (ground truth label) and finds errors in order to modify itself accordingly. Once the algorithm is successfully trained, it will be able to make outcome predictions when applied to new data.
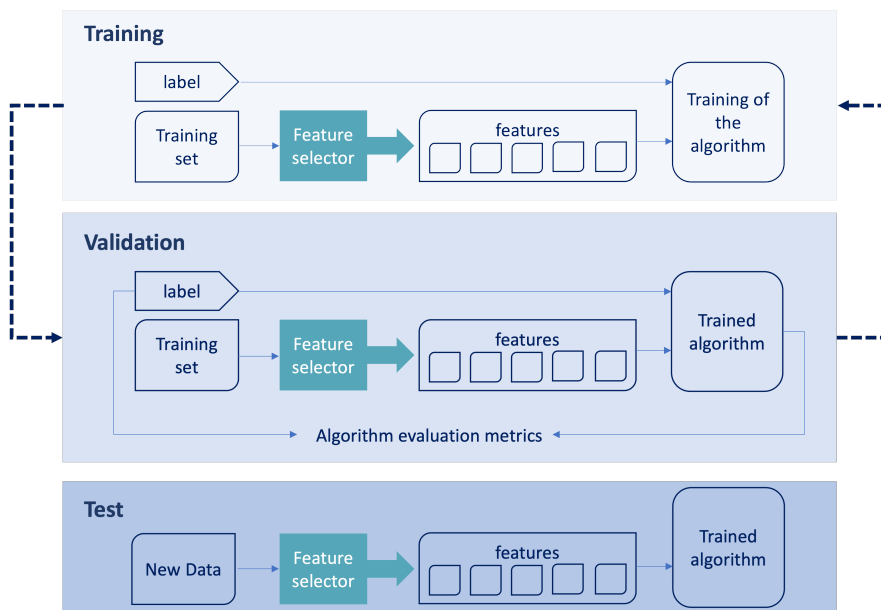
Fig. 3 Three steps of Supervised Learning: i) training of the algorithm, ii) validation of the trained model, iii) test of the model

Predictions can be either discrete (sometimes referred to as classes e.g., positive or negative, benign or malignant, no risk – low risk – high risk, etc.) or continuous (e.g., a value from 0 to 100). A model which maps input to a discrete output is based on a classification algorithm (Fig. 4). Examples of classification algorithms include those which predict if a tumour is benign or malignant, or to establish whether comments written by a patient convey a positive or negative sentiment [15, 16, 17]. Classification algorithms return the probability of a class (between 0 for impossible and 1 for definite). Typically, a probability of 0.50 will be transformed into a class of 1, but this threshold may be modified according to the required

algorithm performance. A model which maps input to a continuous value is based on a regression algorithm (Fig. 4). A regression algorithm might be used, for instance, to estimate the percentage of fat in the liver in case of steatosis, or predict an individual's life expectancy [18, 19].

Classification   Regression
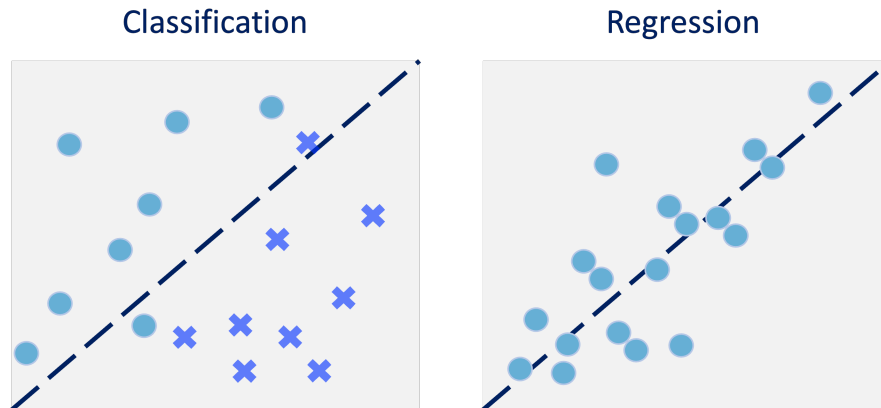
Fig. 4 How classification (on the left) and regression algorithms (on the right) work. Classification algorithms find out the better hyperplane(s) which divides the data in two (or more) classes; a regression model aims to find out the better function that approximates the trend of the data

Note that in [18] the estimation is performed based on ultrasound images. For this type of tasks, i.e. image processing, the predictors must be processed by a feature selector. A feature selector extracts measurable characteristics from the images dataset which then can be represented in a numerical matrix and understood by the algorithm (see Fig. 3).

Four key concerns to be considered in supervised learning are:

- Bias-variance tradeoff: in any supervised model, there is a balance between bias, which is the constant error term, and variance, which is the amount by which the error may vary between different training sets. Increasing bias will usually lead to lower variance, and vice-versa. Generally, in order to produce models that generalize well, the variance of the model should scale with the size and complexity of the training data: small datasets should usually be processed with low-variance models, and large, complex data-sets will often require higher-variance models to fully learn the structure of the data.
- Model complexity and volume of training data: The proper level of model complexity is generally determined by the nature of the training data. A small amount of data, or data that are not uniformly spread throughout different possible scenarios, will be better explained by a low-complexity model. This is because a high-complexity model will overfit if used on a small number of data points.
- Overfitting: it refers to learning a function that fits the training data very well, but does not generalize to other data points. In other words, the model strictly

learns the training data without learning the actual trend or structure in the data themselves that leads to those outputs.

- Noise in the output values: this issue concerns about the amount of noise in the preferred output values. Noisy or incorrect data labels will clearly reduce the effectiveness of the trained model.

Here, the most prominent and common methods used in supervised machine learning are reported: Linear Regression, Support Vector Machine, Random Decision Forest, Extreme Gradient Boosting, Naive Bayes.

Linear regression

Linear Regression (LR) is one of the simplest and most used supervised learning methods. In essence, this method formalizes and identifies the relationship between two or more variables. The assumption of linearity on the cost function is very strong, therefore more complex cost function regression methods have been developed: Non-Linear Regression, Polynomial Regression, Logistic Regression with Sigmoid function, Poisson Regression and many others. As one of the oldest approaches, LR has been widely used in many fields, including medical [20]. This approach is mainly used when a relationship between variables is strongly assumed and the value of one variable (unknown) is to be deduced starting from the values of the other (known). A recent example of use of LR in medicine concerns the prediction of the evolution of systemic diseases starting from clinical evidence [21, 22], or in the genomic field, for example, it can be useful for estimating gene expression patterns in particular biological conditions [23].

Support Vector Machine

Support Vector Machines (SVMs) are supervised learning methods for binary classification. The SVMs represent the data as points in space, building an hyperplane, as wide as possible, which can be positioned as a separator between the two classification categories. The SVMs perform a linear classification, but it is also possible to perform a non-linear classification using an adequate kernel, projecting the data into a multi-dimensional rather than a two-dimensional space. This algorithm is used in many classification and regression problems, in the medical field it is often used for signal separation or for clinical discrimination starting from well-specified characteristics [24, 25]. A very interesting and promising use concerns the early diagnosis [26] or the classification of some types of cancer starting from genomic data [27, 28].

Random Decision Forest

Random Decision Forests (RDF) were first proposed by Tin Kam Ho in 1995 [29]. They are a learning method based on training many Decision Trees (DT), from which a decision strategy is then aggregated. The various Decision Trees are based on the observation of certain characteristics of the data, selected randomly. RDTs are often

unstable methods, but have the great advantage of being easily interpretable. They can be used for both regression and classification. It is mainly used in problems on which there is not yet a precise idea of the weight of the data characteristics or of the relationships between them. In the medical field it is a widely used method for relating clinical features and pathologies [30, 31, 32, 33, 34]. In a recent work by Wang et al. [35] RDFs have been use to detect the factors that most impact on medical expenses for diabetic individuals in the US, while Hsich [36] studied the most critical risk factors for survival in patients with systolic heart failure through Random Survival Forests (RSFs). RDFs models have also been used for the analysis of genomic data [37].

Extreme Gradient Boosting

Extreme Gradient Boosting is a supervised machine learning technique for regression and classification problems that aggregates an ensemble of weak individual models to obtain a more accurate final model. This method is applied to multicollinearity problems in which there are high correlations between the variables. It helps a lot in improving the predictive accuracy of the model and is often used in risk assessment problems. In the medical field it has been used in many applications, some examples are represented by the evaluation of the outcomes of clinical treatments [38] or the study of systemic diseases that depend on multifactorial conditions that are difficult to interpret [39].

Naive Bayes

Bayes Classifiers are ML methods that use Bayes' theorem for the classification process. These classifiers are very fast and, despite their simplicity, are efficient at many complex tasks, even with small training datasets. They are used to calculate the conditional probability of an event, based on the information available on other related events. A disadvantage of such classifiers is the fact that they require knowledge of all the data of the problem, especially the simple and conditional probabilities (information that is difficult to obtain). They also assume the independence of the characteristics of the input and therefore provide a simple approximation (naive) of the problem. In the medical field they have often been used in classification problems [40] or in feature selection [41]. Silla et al. [42] used an extension of the Naive Bayes approach in the context of proteomics, for the hierarchical classification of protein function, while Sandberg et al. [43] used a naive Bayes classifier for the analysis of complete sequences of bacterial genomes, capturing highly specific genomic signatures.

Unsupervised Machine Learning

In contrast with supervised learning, unsupervised learning models process unlabeled data to uncover the underlying data structure. In unsupervised learning,

patterns are found out by algorithms without any input from the user. Unsupervised techniques are thus used to find undefined patterns or clusters of data points which are "closer" or more similar to each other.

A visual illustration of an unsupervised dimension reduction technique is given in Fig. 5. In this figure, the raw data (represented by various shapes in the left panel) are presented to the algorithm which then groups the data into clusters of similar data points (represented in the right panel). Note that data which do not have sufficient commonality to the clustered data are typicaly excluded, thereby reducing the number of features within of the dataset. Indeed, these techniques are often referred to as dimension reduction techniques.
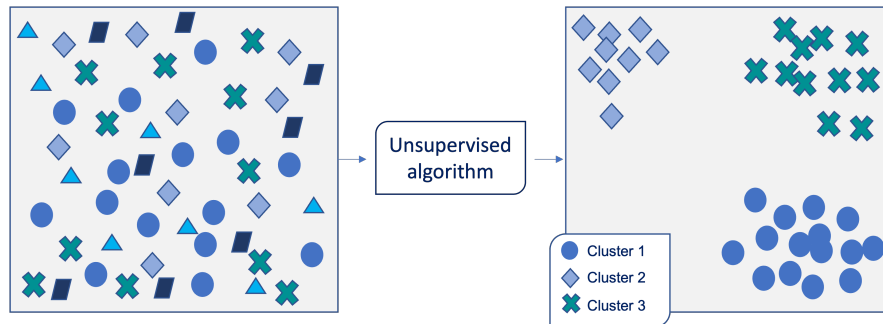
Fig. 5 How Unsupervised Machine Learning algorithms work. They use a more self-contained approach, in which a computer learns to identify patterns without any guidance, only inputting data that is unlabeled and for which no specific output has been defined. In practice, this type of algorithms will learn to divide the data into different clusters based on the characteristics that unite or discriminate them the most.

Unsupervised methods ability to discover similarities and differences in information make them the ideal solution for exploratory data analysis. The output is highly dependent on the algorithm and hyperparameters selected. Hyperparameters, also called tuning parameters, are values used to control the behaviour of the ML algorithm (e.g., number of clusters, distance or density thresholds, type of linkage between clusters). Algorithms exist to detect clusters based on spatial distance between data points, space or subspace density, network connectivity between data points, etc.

By compressing the information in a dataset into fewer features, or dimensions, issues including multiple collinearity between data or high computational cost of the algorithm may be avoided.

Unsupervised approaches also share many similarities to statistical techniques which will be familiar to medical researchers. Unsupervised learning techniques make use of similar algorithms used for clustering and dimension reduction in traditional statistics. Those familiar with Principal Component Analysis, for instance, will already be familiar with many of the techniques used in unsupervised learning.

Here, the most prominent and common methods used in Unsupervised Machine Learning are reported: k-Nearest Neighbours, Principal Component Analysis, k-Means Clustering.

k-Nearest Neighbours

k-Nearest Neighbours is an instance-based learning algorithm, used for pattern recognition for classification or regression. The algorithm uses similarity criteria between elements that are close to each other. In essence the closest neighbours contribute more to the attribution of characteristics than the distant ones. The parameter k represents the number of neighbours that will contribute to the training in the feature space. This algorithm is often used in problems concerning the recognition of similarity patterns aimed at classification and it is a fast, precise and efficient algorithm, but has the disadvantage that the precision of the predictions is strongly dependent on the quality of the data. In the medical field it has often been used for the analysis of hidden patterns on very big amount of data from clinical repositories [44] and in the genomic field [45].

Principal Component Analysis

Principal Component Analysis (PCA, also called Karhunen-Loève transform) is a statistical procedure for dimensionality reduction of the space of variables. The PCA consists of a linear transformation of the variables that projects the original ones into a new Cartesian system in which the new variables try to transfer most of the significance (variance) of the old variables onto a plane, thus obtaining a dimensional reduction without losing too much information. One major limitation of this method is that it can only capture linear correlations between variables. To overcome this disadvantage, sparse PCA and non-linear PCA have been recently introduced. PCA is widely used especially in the fields of medicine and psychology, where scientists works with datasets made up of numerous variables [46, 47, 48].

k-Means Clustering

K-Means Clustering is a vector quantization method for partitioning input data into k clusters. The goal of the algorithm is to minimize the total intra-group variance; each group is identified by a centroid or mid-point. At each step of the algorithm, the input points are assigned to the group with the closest centroid to them. The centroids are recalculated at each step until the algorithm converges and the centroids are stable. k-Means is both a simple and efficient algorithm for clustering problems, but it has the drawback of being very sensitive to outliers, which can significantly deviate the centroids, and of having to choose the number of clusters a priori. In medicine, it is mainly used in situations where a lot of unlabelled data are available [49, 50, 51].

Artificial Neural Networks

Artificial Neural Networks deserve a very long description, which is beyond the scope of this chapter. We will only say that they are computational learning models made up of artificial "neurons", inspired by the simplification of biological neural networks. Such models consist of layers of artificial neurons and processes using computational connections. They are adaptive systems, which change their structure based on the information they process during the learning phase. The layers of neurons can also be very deep (Deep Learning). They are used to create very complex learning algorithms, with very high abstraction capabilities, which are therefore difficult to interpret. In the medical field they have many applications, especially in areas where large amounts of data are available (Big Data) [52, 53, 54, 55]. Among the ANNs, the Convolutional Neural Networks (CNNs) deserve particular attention, in which the pattern of connectivity between neurons is inspired by the organization of the animal visual cortex, for this reason they are particularly suitable for processing images, widely used for the analysis of medical images for detection, segmentation and classification of anomalies or lesions [56, 57, 58].

Reinforcement Learning

Reinforcement learning is a machine learning technique in which a computer (agent) keeps learning continuously to perform a task through repeated trial-and-error interactions with an interactive environment. In other word, agent is self-trained on reward and punishment mechanisms (see Fig. 6). This learning approach allows the agent to make a series of decisions that maximize a reward metric for the activity, without being explicitly programmed to do so and without human intervention.

State
S_t

Action
A_t

Agent

Environment

Reward
R_t
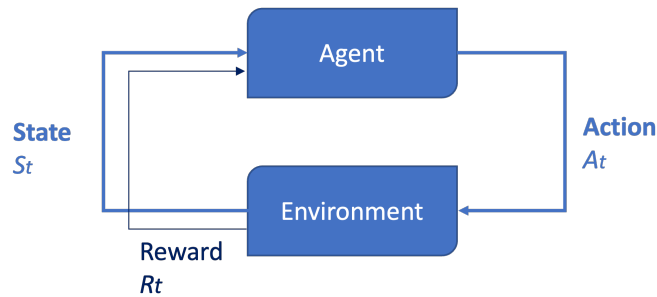
Fig. 6 Basic diagram of Reinforcement Learning

4 Issues and Challenges

Data Management

Data used in ML applications, derived by medical protocols and experimental trials may be incomplete, contain errors, biases and artifacts. Moreover, some data point may be missing for some of the samples. In this scenario, data imputation, denoising and integration should be part of the design of ML algorithms applied to medicine. We will not go into detail, but it is important to know that the performance of AI algorithms is strongly dependent on the quality of the dataset, and that in the case of unbalanced, incomplete, noisy, biased datasets, there are many solutions to be applied to improve datasets and performances [87, 88, 89, 90, 91].

Machine Learning Model Evaluation Metrics

Evaluating the developed machine learning model is an essential part of any project. Here are listed the most widely used evaluation metrics:

- Classification Accuracy: it is the sum of all the correct predictions divided by the total number of input samples. Nevertheless, it works well only in case of balanced dataset, i.e. there are equal number of samples belonging to each class. In case of unbalanced samples, it can be misleading and give the false impression of achieving high accuracy. For instance, if we have 98% samples of class "A" and 2% samples of class "B", the model can easily get 98% training accuracy by simply predicting every training sample belonging to class A. But when tested on a test set with 60% samples of class A and 40% samples of class B, the model can get a lower accuracy of 60%. This issue may be a real problem when the cost of misclassification of the minor class samples is very high: in case of serious pathologies, the cost of not diagnosing a sick person's disease is much higher than the cost of further testing a healthy person.
- Logarithmic Loss (or Log Loss): it penalize the false classifications and works well for multi-class classification. Log Loss nearer to 0 indicates higher accuracy. In general, minimising Log Loss gives higher classifier accuracy.
- Area Under Curve (AUC): it is one of the most widely used metrics for evaluation, especially for binary classification problem. AUC of a classifier indicates the probability that the classifier will rank a randomly chosen positive example higher than a randomly chosen negative example. To understand how it is computed, let us introduce: (i) the True Positive Rate (TPR, or Sensitivity), which corresponds to the proportion of positive data samples that are correctly considered as positive, with respect to all positive data samples; (ii) the True Negative Rate (TNR, or Specificity), which corresponds to the proportion of negative data samples that are correctly considered as negative, with respect to all negative data samples. Sensitivity and (1- Specificity) are plotted at varying threshold values in the range [0,1] and the Receiver Operating Characteristic (ROC) curve is graphed. AUC is

the area under ROC curve and the higher the value, the better is the performance of the ML model.

- Mean Absolute Error: it is the average of the difference between the correct outputs and the predicted outputs. It gives us the measure of how far the predictions were from the actual values. Nevertheless, they don't gives any idea whether we are under predicting the data or over predicting the data.
- Mean Squared Error (MSE): it is quite similar to Mean Absolute Error, as it takes the average of the square of the difference between the correct outputs and the predicted outputs. As the square of the error is calculated, the effect of larger errors become more pronounced then smaller error, hence the model can focus more on larger errors.

Explainability, Interpretability, Ethical and Legal Issues

Both the interpretability and the explainability of ML algorithms go in the direction of strengthening trust in ML algorithms [92, 93]. Interpretability concerns the way in which the ML model reaches its conclusions and has the aim of verifying that the accuracy of the model derives from a correct representation of the problem and not from artifacts present in the training data, while explainability seeks to explain the reasons why an algorithm made one decision rather than another. The concept of explainability is receiving more and more attention from a legal point of view, according to the right of every individual to receive detailed explanations regarding the decisions that impact on his life. In general, explainability and interpretability are well achievable with simple classification algorithms such as decision trees or linear regressors, but the relationships between data characteristics are not always linear, consequently, these approaches may not be suitable in many tasks.

Most ML algorithms are too complex to be understood directly, so it is necessary to adopt post-hoc analysis [94]. In some cases, the explanation is achieved by deriving a transparent model that mimics the original one, although such an approach is not always possible. It is often necessary to work hard to obtain explanations, through empirical approaches, recursive attempts and series of examples. Applications of ML in the medical field are often directed towards diagnosis and therapy, and as with doctors, they can make mistakes, generate delays and sometimes lead towards wrong therapies, running into major legal and economic problems. For this reason, the use of ML in the medical field is still limited to supporting diagnosis, providing for the review of outputs by human experts.

Another very important aspect of the application of ML in medicine concerns the equity of data. By their nature, and certainly not by intention, artificial intelligence algorithms can be highly discriminatory against minorities, generating deep ethical issues [95, 96]. Given their heavy dependence on available data, it is clear that their performance will be very good for people from well-represented populations and very bad, misleading and even dangerously wrong for people from underrepresented populations [97]. With most medical data repositories coming from hospitals in rich, industrialized countries, developing countries risk finding themselves completely

shut out of new medical advances if they rely heavily on ML. This is certainly a problem that needs attention and prompt solutions.

Perspectives in Personalised Medicine

The synergy between artificial intelligence and precision medicine is revolutionising the concepts of diagnosis, prognosis and assistance [98, 99]. Conventional symptom-based treatment of patients is slowly giving way to more holistic approaches in which aggregate data and biological indicators of each patient are combined with specific observations and general patterns inferred from artificial intelligence approaches on large numbers of patients . Genetics, genomics, precision medicine, combined with machine learning and deep learning algorithms, make it possible to generate a personalized therapy even for individuals with less common therapeutic responses or with particular medical needs [100, 101, 102].

In general, the shared adoption of the EHR (Electronic Health Record) format for storing clinical data has allowed for the systematic collection of a great deal of information on the health of individuals in digital format. This format has greatly facilitated the use of ML tools in the medical field, thanks to the uniformity of data and the ease of retrieval and use. Other projects born with the aim of facilitating the application of ML in the medical field are the health databases containing data of millions of individuals, such as the All of US Research Program Project, the Human Genome Project, the UK Biobank, the IARC Biobank and the European Biobank.

To conclude our contribute on the innovative, indeed, revolutionary perspectives of the application of ML in medicine, we would like to mention the Digital Twins, models that mimic the biological background of a patient as much as possible, making it possible to test drugs, therapies, treatments, maximizing the results and minimizing the risks on the patient's health [103].

5 Conclusions

The introduction of Artificial Intelligence has upset all fields of research, necessarily medicine as well. It has changed our perception of diseases, of treatments, our relationship with doctors and it has changed procedures, it has opened many doors and brought to the table issues and problems that we hadn't really thought about yet. We consider ML as a huge opportunity to improve more or less everything we can operate on, but it is a tool to understand, to handle with care, to use with attention. In this chapter we have provided some basic indications to navigate the ocean of ML, especially for non-experts. We've described the most used methods, giving guidance on when to use them, and we've provided plenty of bibliography, so you know where to go further. We have tried not so much to provide answers, but rather to help understand what are the right questions to ask when you want to use ML in the medical field, which is already a very important starting point.

USEFUL GLOSSARY

- **Accuracy:** Measure of the algorithm ability in giving correct predictions.
- **Algorithm:** Any systematic calculation scheme or procedure.
- **Classification:** Learning process in which the data are divided into two or more classes and the system assigns one or more classes among the available ones to an input.
- **Clustering:** Learning process in which a set of data is divided into groups that are not known a priori.
- **Features:** Interesting parts, qualities or characteristics of something.
- **Layer:** Collection of nodes operating together.
- **Model:** Formal representation of knowledge related to a phenomenon.
- **Normalisation:** Process of feature transformation to obtain a similar scale.
- **Neural Networks:** Computational model made of artificial neurons, vaguely inspired by the simplification of a biological neural network.
- **Node:** Computational unit (also called artificial neuron) which receives one or more inputs and composes them to produce an output.
- **Overfitting:** Excessive reliance of the model on training data, leading to inability to generalise and evaluate other data well.
- **Pre-processing:** Manipulating data before it is used in order to ensure or improve performance in the data mining process.
- **Regression:** Learning process similar to classification, with the difference that the output has a continuous domain and not a discrete one.
- **Training:** the process of creating a model from the training data. The data is fed into the training algorithm, which learns a representation for the problem, and produces a model. Also called "learning".
- **Training Set:** Set of data used in input during the learning process to fit the parameters.
- **Test Set:** Set of data, independent to the training set, used only to assess the performances of a fully specified classifier or a regressor.
- **Validation Set:** Set of data used to tune the parameters and to assess the performances of a classifier or a regressor. It is sometimes also called the development set (dev set).
- **Weights:** Parameters within a neural network calibrating the transformation of input data.

References

1. Wartman, S. Combs, C. Medical Education Must Move From the Information Age to the Age of Artificial Intelligence. *Academic Medicine*. **93** pp. 1107-1109 (2018)

2. Obermeyer, Z. Lee, T. Lost in Thought - The Limits of the Human Mind and the Future of Medicine.. *The New England Journal Of Medicine*. **377 13** pp. 1209-1211 (2017)

3. Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M., Zhao, S. Applications of machine learning in drug discovery and development. *Nat Rev Drug Discov*. **18** pp.463-477 (2019)

4. Valletta, J.J., Torney, C., Kings, M., Thornton, A., Madden, J. Applications of machine learning in animal behaviour studies *Animal Behaviour*. **124** pp.203-220 (2017)

5. Recknagel, F. Applications of machine learning to ecological modelling. *Ecological Modelling*. **146** pp.303-310 (2001)

6. Garg, A. Mago, V. Role of machine learning in medical research: A survey. *Comput. Sci. Rev.*. **40** pp. 100370 (2021)

7. Rajkomar, A., Dean, J. Kohane, I. Machine Learning in Medicine. *The New England Journal Of Medicine*. **380** pp. 1347-1358 (2019)

8. Erickson, B., Korfiatis, P., Akkus, Z. Kline, T. Machine Learning for Medical Imaging.. *Radiographics : A Review Publication Of The Radiological Society Of North America, Inc*. **37 2** pp. 505-515 (2017)

9. Frankish, K. Ramsey, W. The Cambridge Handbook of Artificial Intelligence. (2014)

10. Smiti, A. When machine learning meets medical world: Current status and future challenges *Computer Science Review*. **37** (2020)

11. Garg, A., Mago, V. Role of machine learning in medical research: A survey *Computer Science Review*. **40** (2021)

12. Brynjolfsson, E. McAfee, A. The second machine age: work, progress, and prosperity in a time of brilliant technologies, 1st Edition. (2014)

13. Susskind, R. Susskind, D. The Future of the Professions: How Technology Will Transform the Work of Human Experts. (2016)

14. Parish, J. The Patient Will See You Now: The Future of Medicine is in Your Hands.. *Journal Of Clinical Sleep Medicine : JCSM : Official Publication Of The American Academy Of Sleep Medicine*. (2015)

15. Esteva, A., Kuprel, B., Novoa, R., Ko, J., Swetter, S., Blau, H. Thrun, S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. **542** pp. 115-118 (2017)

16. Hawkins, J., Brownstein, J., Tuli, G., Runels, T., Broecker, K., Nsoesie, E., McIver, D., Rozenblum, R., Wright, A., Bourgeois, F. Greaves, F. Measuring patient-perceived quality of care in US hospitals using Twitter. *BMJ Quality Safety*. **25** pp. 404 - 413 (2015)

17. Mangasarian, O., Street, W. Wolberg, W. Breast Cancer Diagnosis and Prognosis Via Linear Programming. *Oper. Res.*. **43** pp. 570-577 (1995)

18. Colantonio, S., Salvati, A., Caudai, C., Bonino, F., De Rosa, L., Pascali, M. A., Germanese, D., Brunetto, M. R., Faita, F. A deep learning approach for hepatic steatosis estimation from ultrasound imaging. *Proc. of ICCCI 2021 - 13th International Conference on Computational Collective Intelligence, Rhodes, Greece* pp. 703–714 (2021)

19. Ali, N., Srivastava, D., Tiwari, A., Pandey, A., Pandey, A. K., Sahu, A. Predicting Life Expectancy of Hepatitis B Patients using Machine Learning. *IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*. (2022)

20. Sidey-Gibbons, J. Sidey-Gibbons, C. Machine learning in medicine: a practical introduction. *BMC Medical Research Methodology*. **19** (2019)

21. Simos, N., Manikis, G., Papadaki, E., Kavroulakis, E., Bertsias, G. Marias, K. Machine Learning Classification of Neuropsychiatric Systemic Lupus Erythematosus patients using resting-state fMRI functional connectivity. *2019 IEEE International Conference On Imaging Systems And Techniques (IST)*. pp. 1-6 (2019)

22. Coelewij, L., Waddington, K., Robinson, G., Chocano, E., McDonnell, T., Farinha, F., Peng, J., Dönnes, P., Smith, E., Croca, S., Bakshi, J., Griffin, M., Nicolaides, A., Rahman, A., Jury, E. Pineda-Torra, I. Serum Metabolomic Signatures Can Predict Subclinical Atherosclerosis in Patients With Systemic Lupus Erythematosus.. *Arteriosclerosis, Thrombosis, Vascular Biology*. (2021)

23. Liu, S., Lu, M., Li, H. Zuo, Y. Prediction of Gene Expression Patterns With Generalized Linear Regression Model. *Frontiers In Genetics*. **10** (2019)

24. Taylor, R., Moore, C., Cheung, K. Brandt, C. Predicting urinary tract infections in the emergency department with machine learning. *PLoS ONE*. **13** (2018)

25. Leha, A., Hellenkamp, K., Unsoeld, B., Mushemi-Blake, S., Shah, A., Hasenfuss, G. Seidler, T. A machine learning approach for the prediction of pulmonary hypertension. *PLoS ONE*. **14** (2019)

26. Zhang, F., Kaufman, H., Deng, Y. Drabier, R. Recursive SVM biomarker selection for early detection of breast cancer in peripheral blood. *BMC Medical Genomics*. **6** pp. S4 - S4 (2013)

27. Huang, S., Cai, N., Pacheco, P., Narrandes, S., Wang, Y. Xu, W. Applications of Support Vector Machine (SVM) Learning in Cancer Genomics.. *Cancer Genomics Proteomics*. **15 1** pp. 41-51 (2018)

28. Liao, C. Li, S. A Support Vector Machine Ensemble for Cancer Classification Using Gene Expression Data. *ISBRA*. (2007)

29. Ho, T. Random decision forests. *Proceedings Of 3rd International Conference On Document Analysis And Recognition*. **1** pp. 278-282 vol.1 (1995)

30. Zhu, M., Xia, J., Jin, X., Yan, M., Cai, G., Yan, J. Ning, G. Class Weights Random Forest Algorithm for Processing Class Imbalanced Medical Data. *IEEE Access*. **6** pp. 4641-4652 (2018)

31. Martin-Gutierrez, L., Peng, J., Thompson, N., Robinson, G., Naja, M., Peckham, H., Wu, W., J'bari, H., Ahwireng, N., Waddington, K., Bradford, C., Varnier, G., Gandhi, A., Radmore, R., Gupta, V., Isenberg, D., Jury, E. Ciurtin, C. Two shared immune cell signatures stratify patients with Sjögren's syndrome and systemic lupus erythematosus with potential therapeutic implications.. *Arthritis Rheumatology*. (2021)

32. Seccia, R., Gammelli, D., Dominici, F., Romano, S., Landi, A., Salvetti, M., Tacchella, A., Zaccaria, A., Crisanti, A., Grassi, F. Palagi, L. Considering patient clinical history impacts performance of machine learning models in predicting course of multiple sclerosis. *PLoS ONE*. **15** (2020)

33. Vlahou, A., Schorge, J., Gregory, B. Coleman, R. Diagnosis of Ovarian Cancer Using Decision Tree Classification of Mass Spectral Data. *Journal Of Biomedicine And Biotechnology*. **2003** pp. 308 - 314 (2003)

34. Baumgartner, C., Böhm, C. Baumgartner, D. Modelling of classification rules on metabolic patterns including machine learning and expert knowledge. *Journal Of Biomedical Informatics*. **38 2** pp. 89-98 (2005)

35. Wang, J. Shi, L. Prediction of medical expenditures of diagnosed diabetics and the assessment of its related factors using a random forest model, MEPS 2000-2015.. *International Journal For Quality In Health Care : Journal Of The International Society For Quality In Health Care*. (2020)

36. Hsich, E., Gorodeski, E., Blackstone, E., Ishwaran, H. Lauer, M. Identifying Important Risk Factors for Survival in Patient With Systolic Heart Failure Using Random Survival Forests. *Circulation: Cardiovascular Quality And Outcomes*. **4** pp. 39-45 (2011)

37. Chen, X. Ishwaran, H. Random forests for genomic data analysis.. *Genomics*. **99 6** pp. 323-9 (2012)

38. Mo, X., Chen, X., Ieong, C., Zhang, S., Li, H., Li, J., Lin, G., Guang-Sun, He, F., Ya-He, Xie, Y., Zeng, P., Chen, Y., Liang, H. Zeng, H. Early Prediction of Clinical Response to Etanercept Treatment in Juvenile Idiopathic Arthritis Using Machine Learning. *Frontiers In Pharmacology*. **11** (2020)

39. Murray, S., Avati, A., Schmajuk, G. Yazdany, J. Automated and flexible identification of complex disease: building a model for systemic lupus erythematosus using noisy labeling. *Journal Of The American Medical Informatics Association : JAMIA*. **26 1** pp. 61-65 (2019)

40. D'souza, K. Ansari, Z. Big Data Science in Building Medical Data Classifier Using Naïve Bayes Model. *2018 IEEE International Conference On Cloud Computing In Emerging Markets (CCEM)*. pp. 76-80 (2018)

41. Degroeve, S., Baets, B., Peer, Y. & Rouzé, P. Feature subset selection for splice site prediction. *Bioinformatics*. **18 Suppl 2** pp. S75-83 (2002)

42. Silla, C. Freitas, A. A Global-Model Naive Bayes Approach to the Hierarchical Prediction of Protein Functions. *2009 Ninth IEEE International Conference On Data Mining*. pp. 992-997 (2009)

43. Sandberg, R., Winberg, G., Brändén, C., Kaske, A., Ernberg, I. Cöster, J. Capturing whole-genome characteristics in short sequences using a naïve Bayesian classifier.. *Genome Research*. **11 8** pp. 1404-9 (2001)

44. Khamis, H. Application of k-Nearest Neighbour classification in medical data in the context of Kenia. *Digital Repository Unimib*. (2014)

45. Parry, R., Jones, W., Stokes, T., Phan, J., Moffitt, R., Fang, H., Shi, L., Oberthuer, A., Fischer, M., Tong, W. Wang, M. k-Nearest neighbor models for microarray gene expression analysis and clinical outcome prediction. *The Pharmacogenomics Journal*. **10** pp. 292 - 309 (2010)

46. Alexe, G., Dalgin, G., Ganesan, S., DeLisi, C. Bhanot, G. Analysis of breast cancer progression using principal component analysis and clustering. *Journal Of Biosciences*. **32** pp. 1027-1039 (2007)

47. Maisuradze, G., Liwo, A. Scheraga, H. Principal component analysis for protein folding dynamics.. *Journal Of Molecular Biology*. **385 1** pp. 312-29 (2009)

48. Taguchi, Y. Okamoto, A. Principal component analysis for bacterial proteomic analysis. *2011 IEEE International Conference On Bioinformatics And Biomedicine Workshops (BIBMW)*. pp. 961-963 (2011)

49. Le, T. Fuzzy C-Means Clustering Interval Type-2 Cerebellar Model Articulation Neural Network for Medical Data Classification. *IEEE Access*. **7** pp. 20967-20973 (2019)

50. Khanmohammadi, S., Adibeig, N. Shanehbandy, S. An improved overlapping k-means clustering method for medical applications. *Expert Syst. Appl.*. **67** pp. 12-18 (2017)

51. Handhayani, T. Hiryanto, L. Intelligent Kernel K-Means for Clustering Gene Expression. *Procedia Computer Science*. **59** pp. 171-177 (2015)

52. Greenspan, H., Ginneken, B. Summers, R. Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique. *IEEE Trans. Medical Imaging*. **35** pp. 1153-1159 (2016)

53. Litjens, G., Kooi, T., Bejnordi, B., Setio, A., Ciompi, F., Ghafoorian, M., Laak, J., Ginneken, B. Sánchez, C. A survey on deep learning in medical image analysis. *Medical Image Analysis*. **42** pp. 60-88 (2017)

54. Gao, X., Lin, S. Wong, T. Automatic Feature Learning to Grade Nuclear Cataracts Based on Deep Learning. *IEEE Transactions On Biomedical Engineering*. **62** pp. 2693-2701 (2015)

55. Sundaram, L., Gao, H., Padigepati, S., McRae, J., Li, Y., Kosmicki, J., Fritzilas, N., Hakenberg, J., Dutta, A., Shon, J., Xu, J., Batzoglou, S., Li, X. Farh, K. Predicting the clinical impact of human mutation with deep neural networks. *Nature Genetics*. **50** pp. 1161 - 1170 (2018)

56. Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J. Greenspan, H. GAN-based Synthetic Medical Image Augmentation for increased CNN Performance in Liver Lesion Classification. *Neurocomputing*. **321** pp. 321-331 (2018)

57. Setio, A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., Riel, S., Wille, M., Naqibullah, M., Sánchez, C. Ginneken, B. Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks. *IEEE Transactions On Medical Imaging*. **35** pp. 1160-1169 (2016)

58. Kamnitsas, K., Ledig, C., Newcombe, V., Simpson, J., Kane, A., Menon, D., Rueckert, D. Glocker, B. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis*. **36** pp. 61-78 (2017)

59. Poplin, R., Varadarajan, A., Blumer, K., Liu, Y., McConnell, M., Corrado, G., Peng, L. Webster, D. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nature Biomedical Engineering*. **2** pp. 158-164 (2018)

60. Rumsfeld, J., Joynt, K. Maddox, T. Big data analytics to improve cardiovascular care: promise and challenges. *Nature Reviews Cardiology*. **13** pp. 350-359 (2016)

61. Ramalingam, V., Dandapath, A. Raja, M. Heart disease prediction using machine learning techniques : a survey. *International Journal Of Engineering And Technology*. **7** pp. 684 (2018)

62. Zhou, M., Scott, J., Chaudhury, B., Hall, L., Goldgof, D., Yeom, K., Iv, M., Ou, Y., Kalpathy-Cramer, J., Napel, S., Gillies, R., Gevaert, O. Gatenby, R. Radiomics in Brain Tumor: Image Assessment, Quantitative Feature Descriptors, and Machine-Learning Approaches. *American Journal Of Neuroradiology*. **39** pp. 208 - 216 (2018)

63. Koh, D., Papanikolaou, N., Bick, U., Illing, R., Kahn, C., Kalpathi-Cramer, J., Matos, C., Martí-Bonmatí, L., Miles, A., Mun, S., Napel, S., Rockall, A., Sala, E., Strickland, N. Prior, F. Artificial intelligence and machine learning in cancer imaging. *Communications Medicine*. **2** (2022)

64. Zerouaoui, H. Idri, A. Reviewing Machine Learning and Image Processing Based Decision-Making Systems for Breast Cancer Imaging. *Journal Of Medical Systems*. **45** pp. 1-20 (2021)

65. Bi, W., Hosny, A., Schabath, M., Giger, M., Birkbak, N., Mehrtash, A., Allison, T., Arnaout, O., Abbosh, C., Dunn, I., Mak, R., Tamimi, R., Tempany, C., Swanton, C., Hoffmann, U., Schwartz, L., Gillies, R., Huang, R. Aerts, H. Artificial intelligence in cancer imaging: Clinical challenges and applications. *Ca*. **69** pp. 127 - 157 (2019)

66. Kircher, M., Witten, D., Jain, P., O'Roak, B., Cooper, G. Shendure, J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*. **46** pp. 310 - 315 (2014)

67. Hsu, Y., Huang, P. Chen, D. Sparse principal component analysis in cancer research.. *Translational Cancer Research*. **3** 3 pp. 182-190 (2014)

68. Chen, L., Li, H., Xie, L., Zuo, Z., Tian, L., Liu, C. Guo, X. Editorial: Big Data and Machine Learning in Cancer Genomics. *Frontiers In Genetics*. **12** (2021)

69. Pan, X. Multiple linear regression for protein secondary structure prediction. *Proteins: Structure*. **43** (2001)

70. Cao, C., Liu, F., Tan, H., Song, D., Shu, W., Li, W., Zhou, Y., Bo, X. Xie, Z. Deep Learning and Its Applications in Biomedicine. *Genomics, Proteomics Bioinformatics*. **16** pp. 17 - 32 (2018)

71. Asgari, E., Mofrad, M. Kobeissy, F. Continuous Distributed Representation of Biological Sequences for Deep Proteomics and Genomics. *PLoS ONE*. **10** (2015)

72. Mathlin, J., Pera, L. Colombo, T. A Census and Categorization Method of Epitranscriptomic Marks. *International Journal Of Molecular Sciences*. **21** (2020)

73. Caudai, C., Galizia, A., Geraci, F., Pera, L., Morea, V., Salerno, E., Via, A. Colombo, T. AI applications in functional genomics. *Computational And Structural Biotechnology Journal*. **19** pp. 5762 - 5790 (2021)

74. Da-Yuan, Liang, Y., Yi, L., Xu, Q. Kvalheim, O. Uncorrelated linear discriminant analysis (ULDA): A powerful tool for exploration of metabolomics data. *Chemometrics And Intelligent Laboratory Systems*. **93** pp. 70-79 (2008)

75. Alakwaa, F., Chaudhary, K. Garmire, L. Deep Learning Accurately Predicts Estrogen Receptor Status in Breast Cancer Metabolomics Data. *Journal Of Proteome Research*. **17** pp. 337 - 347 (2017)

76. Khodayari, A. Maranas, C. A genome-scale Escherichia coli kinetic metabolic model k-ecoli457 satisfying flux data for multiple mutant strains. *Nature Communications*. **7** (2016)

77. Yang, H., Yu, B., Ouyang, P., Li, X., Lai, X., Zhang, G. Zhang, H. Machine learning-aided risk prediction for metabolic syndrome based on 3 years study. *Scientific Reports*. **12** (2022)

78. Grinfeld, J., Nangalia, J., Baxter, J., Wedge, D., Angelopoulos, N., Cantrill, R., Godfrey, A., Papaemmanuil, E., Gundem, G., Maclean, C., Cook, J., O'Neil, L., O'meara, S., Teague, J., Butler, A., Massie, C., Williams, N., Nice, F., Andersen, C., Hasselbalch, H., Guglielmelli, P., McMullin, M., Vannucchi, A., Harrison, C., Gerstung, M., Green, A. Campbell, P. Classification and Personalized Prognosis in Myeloproliferative Neoplasms. *The New England Journal Of Medicine*. **379** pp. 1416-1430 (2018)

79. Denis, F., Basch, E., Septans, A., Bennouna, J., Urban, T., Dueck, A. Letellier, C. Two-Year Survival Comparing Web-Based Symptom Monitoring vs Routine Surveillance Following Treatment for Lung Cancer. *JAMA*. **321** pp. 306-307 (2019)

80. Hasnain, Z., Mason, J., Gill, K., Miranda, G., Gill, I., Kuhn, P. Newton, P. Machine learning models for predicting post-cystectomy recurrence and survival in bladder cancer patients. *PLoS ONE*. **14** (2019)

81. Nie, D., Zhang, H., Adeli, E., Liu, L. Shen, D. 3D Deep Learning for Multi-modal Imaging-Guided Survival Time Prediction of Brain Tumor Patients. *Medical Image Computing And Computer-assisted Intervention : MICCAI ... International Conference On Medical Image Computing And Computer-Assisted Intervention*. **9901** pp. 212-220 (2016)

82. Meiring, C., Dixit, A., Harris, S., MacCallum, N., Brealey, D., Watkinson, P., Jones, A., Ashworth, S., Beale, R., Brett, S., Singer, M. Ercole, A. Optimal intensive care outcome prediction over time using machine learning. *PLoS ONE*. **13** (2018)

83. Chen, H., Engkvist, O., Wang, Y., Olivecrona, M. Blaschke, T. The rise of deep learning in drug discovery.. *Drug Discovery Today*. **23 6** pp. 1241-1250 (2018)

84. Gupta, S., Chaudhary, K., Kumar, R., Gautam, A., Nanda, J., Dhanda, S., Brahmachari, S. Raghava, G. Prioritization of anticancer drugs against a cancer using genomic features of cancer cells: A step towards personalized medicine. *Scientific Reports*. **6** (2016)

85. Hejase, H. Chan, C. Improving Drug Sensitivity Prediction Using Different Types of Data. *CPT: Pharmacometrics Systems Pharmacology*. **4** (2015)

86. Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M. Zhao, S. Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery*. **18** pp. 463-477 (2019)

87. Jain, A., Patel, H., Nagalapatti, L., Gupta, N., Mehta, S., Guttula, S., Mujumdar, S., Afzal, S., Mittal, R. Munigala, V. Overview and Importance of Data Quality for Machine Learning Tasks. *Proceedings Of The 26th ACM SIGKDD International Conference On Knowledge Discovery Data Mining*. (2020)

88. Dai, W., Yoshigoe, K. Parsley, W. Improving Data Quality through Deep Learning and Statistical Models. *ArXiv*. **abs/1810.07132** (2018)

89. Luca, A., Ursuleanu, T., Gheorghe, L., Grigorovici, R., Iancu, S., Hlusneac, M. Grigorovici, A. Impact of quality, type and volume of data used by deep learning models in the analysis of medical images. *Informatics In Medicine Unlocked*. (2022)

90. Wang, Z., Poon, J., Sun, S. Poon, S. Attention-based Multi-instance Neural Network for Medical Diagnosis from Incomplete and Low Quality Data. *2019 International Joint Conference On Neural Networks (IJCNN)*. pp. 1-8 (2019)

91. Chang, Y., Yan, L., Chen, M., Fang, H. Zhong, S. Two-Stage Convolutional Neural Network for Medical Noise Removal via Image Decomposition. *IEEE Transactions On Instrumentation And Measurement*. **69** pp. 2707-2721 (2020)

92. Marcinkevics, R. Vogt, J. Interpretability and Explainability: A Machine Learning Zoo Mini-tour. *ArXiv*. **abs/2012.01805** (2020)

93. Samek, W. Müller, K. Towards Explainable Artificial Intelligence. *ArXiv*. **abs/1909.12072** (2019)

94. Montavon, G., Samek, W. Müller, K. Methods for interpreting and understanding deep neural networks. *ArXiv*. **abs/1706.07979** (2018)

95. Chen, I., Pierson, E., Rose, S., Joshi, S., Ferryman, K. Ghassemi, M. Ethical Machine Learning in Health Care. *Annual Review Of Biomedical Data Science*. **4** pp. 123-144 (2021)

96. Yoon, C., Torrance, R. Scheinerman, N. Machine learning in medicine: should the pursuit of enhanced interpretability be abandoned?. *Journal Of Medical Ethics*. **48** pp. 581 - 585 (2021)

97. Martin, A., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. Daly, M. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature Genetics*. **51** pp. 584-591 (2019)

98. Johnson, K., Wei, W., Weeraratne, D., Frisse, M., Misulis, K., Rhee, K., Zhao, J. Snowdon, J. Precision Medicine, AI, and the Future of Personalized Health Care. *Clinical And Translational Science*. **14** pp. 86 - 93 (2020)

99. Quazi, S. Artificial intelligence and machine learning in precision and genomic medicine. *Medical Oncology (Northwood, London, England)*. **39** (2022)

100. Xu, J., Yang, P., Xue, S., Sharma, B., Sanchez-Martin, M., Wang, F., Beaty, K., Dehan, E. Parikh, B. Translating cancer genomics into precision medicine with artificial intelligence: applications, challenges and future perspectives. *Human Genetics*. **138** pp. 109 - 124 (2019)

101. Grapov, D., Fahrmann, J., Wanichthanarak, K. Khoomrung, S. Rise of Deep Learning for Genomic, Proteomic, and Metabolomic Data Integration in Precision Medicine. *OMICS : A Journal Of Integrative Biology*. **22** pp. 630 - 636 (2018)

102. Hamamoto, R., Komatsu, M., Takasawa, K., Asada, K. Kaneko, S. Epigenetics Analysis and Integrated Analysis of Multiomics Data, Including Epigenetic Data, Using Artificial Intelligence in the Era of Precision Medicine. *Biomolecules*. **10** (2019)

103. Björnsson, B., Borrebaeck, C., Elander, N., Gasslander, T., Gawel, D., Gustafsson, M., Jörnsten, R., Lee, E., Li, X., Lilja, S., Martínez-Enguita, D., Matussek, A., Sandström, P., Schäfer, S., Stenmarker, M., Sun, X., Sysoev, O., Zhang, H. Benson, M. Digital twins to personalize medicine. *Genome Medicine*. **12** (2019)