AI: A Glossary of Terms

Disclaimer

The glossary of terms contains entries that we think might come handy when studying *Artificial Intelligence in Medical Imaging*. Many of the terms can be found in the preceding chapters. Some descriptions were found on the Internet (see excellent websites such as Techopedia and Medium). In most cases no authors could be traced. If sources were identified, we obtained permission to reproduce. For legibility we avoided mentioning the sources at each entry. References will be given upon request. The authors appreciate feedback if sources are unrightfully omitted.

Thanks to (in alphabetical order): Ameen Abu-Hanna, Anjum Ahmed, Brad Genereaux, Peter van Ooijen, and Martijn Schut.

Jiapan Guo, Postdoc in Medical Imaging Informatics in University Medical Center Groningen, The Netherlands (j.guo@umcg.nl)

Violet Farhang-Razi, MD Northwest Hospital Alkmaar, The Netherlands (v.farhang-razi@nwz.nl)

Paul Algra MD PhD, neuroradiologist Northwest Hospital Alkmaar, The Netherlands (p.r.algra@nwz.nl)

Assembled by Jiapan Guo, Violet Farhang-Razi and Paul Algra (editor).

E. R. Ranschaert et al. (eds.), *Artificial Intelligence in Medical Imaging*,
https://doi.org/10.1007/978-3-319-94878-2

Glossary

A

Algorithm A formula or set of rules (or procedure, processes, or instructions) for solving a problem or for performing a task. In Artificial Intelligence, the algorithm tells the machine how to find answers to a question or solutions to a problem. In machine learning, systems use many different types of algorithms. Common examples include decision trees, clustering algorithms, classification algorithms, or regression algorithms.

AlexNet The name of a neural network that won the ImageNet Large Scale Visual Recognition Challenge in 2012. It is named after Alex Krizhevsky, then a computer science PhD student at Stanford University. See *ImageNet.*

AlphaGo AlphaGo is the first computer program that defeated a professional player on the board game Go in October 2015. Later in October 2017, AlphaGo's team released its new version named AlphaGo Zero which is stronger than any previous human-champion-defeating versions. Go is played on 19 by 19 board which allows for 10^{171} possible layouts (chess 10^{50} configurations). It is estimated that there are 10^{80} atoms in the universe.

Analogical Reasoning Solving problems by using analogies, by comparing to past experiences.

Anonymization The process in which data is de-identified as part of a mechanism to submit data for machine learning.

Area under curve (AUC) The area under a curve between two points is calculated by performing the definite integral. In the context of a receiver operating characteristic for a binary classifier, the AUC represents the classifier's accuracy.

Artificial Intelligence (AI) Artificial intelligence (or machine intelligence) refers to systems that display intelligent behavior by analyzing their environment and taking actions—with some degree of autonomy—to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g., voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (e.g., advanced robots, autonomous cars, drones, or Internet of Things applications). The term AI was first coined by John McCarthy in 1956.

Artificial Intelligence complete AI-complete, which is short for Artificial Intelligence complete or sometimes called AI-hard, describes the complexity of the computational problems is equal to that of the entire AI problem which aims at producing a general computerized system with the human-level intelligence. An AI-complete problem addresses the fact that the problem cannot be easily solved by a simple specific algorithm.

Artificial Intelligence Winters (AIWI) Artificial Intelligence Winters are periods of time during which artificial intelligence experienced reduced fundings for researches

E. R. Ranschaert et al. (eds.), *Artificial Intelligence in Medical Imaging*,
https://doi.org/10.1007/978-3-319-94878-2

and low interest from the public. Two major winter periods were in 1974–1980 and 1987–1993. AIW are the result of inflated, unreal expectations.

Artificial General Intelligence (AGI) Artificial general intelligence as opposed to narrow intelligence, also known as complete, strong, super intelligence, Human Level Machine Intelligence, indicates the ability of a machine that can successfully perform any tasks in an intellectual way as the human being. Artificial superintelligence is a term referring to the time when the capability of computers will surpass humans.

Artificial Superintelligence (ASI) Artificial superintelligence is a term referring to the time when the capability of computers will surpass humans. "Artificial intelligence," which has been much used since the 1970s, refers to the ability of computers to mimic human thought. Artificial superintelligence goes a step beyond and posits a world in which a computer's cognitive ability is superior to a human's.

Artificial Narrow Intelligence (ANI) Artificial Narrow Intelligence, also known as weak or applied intelligence, represents most of the current artificial intelligent systems which usually focus on a specific task. Narrow AIs are mostly much better than humans at the task they were made for: for example, look at face recognition, chess computers, calculus, and translation. The definition of artificial narrow intelligence is in contrast to that of strong AI or artificial general intelligence, which aims at providing a system with consciousness or the ability to solve any problems. Virtual assistants and AlphaGo are examples of artificial narrow intelligence systems.

Artificial Neural Network (ANN) Artificial Neural Network (ANN) is a computational model in machine learning, which is inspired by the biological structures and functions of the mammalian brain. Such a model consists of multiple units called artificial neurons which build connections between each other to pass information. The advantage of such a model is that it progressively "learns" the tasks from the given data without specific programing for a single task.

Artificial Neuron An artificial neuron is a digital construct that seeks to simulate the behavior of a biological neuron in the brain. Artificial neurons are typically used to make up an artificial neural network—these technologies are modeled after human brain activity.

Asimov Isaac Asimov's Three Laws are as follows: (1) A robot may not injure a human being. (2) A robot must obey orders, unless they conflict with law number one. (3) A robot must protect its own existence, as long as those actions do not conflict with either the first or second law.

Association Subcategory of unsupervised learning. It can be best explained by market basket analysis (MBA). MBA attempts to identify association/relation between various items that have been chosen by a particular shopper and placed in their respective baskets (real or virtual). The output value from this lies in cross marketing of products and customer behavior analysis. Association is the generalization of m.b.a. Example: there is a good chance that a customer will buy bread if he has already bought milk and eggs.

Augmented Intelligence Augmented Intelligence is the intersection of machine learning and advanced applications, where clinical knowledge and medical data converge on a single platform. The potential benefits of Augmented Intelligence are realized when it is used in the context of workflows and systems that healthcare practitioners operate and interact with. Unlike Artificial Intelligence, which tries to replicate human intelligence, Augmented Intelligence works with and amplifies human intelligence.

Autoregressive Model An autoregressive model is a time series model that uses observations from previous time steps as input to a regression equation to predict the value at the next time step. In statistics and signal processing, an autoregressive model is a representation of a type of random process. It is used to describe certain time-varying processes in nature, economics, etc.

B

Backpropagation Backpropagation, also called "backward propagation of errors," is an approach that is commonly used in the training process of the deep neural network to reduce errors. It allows the machine learning algorithm to adjust itself according to looking at its past function. It involves the calculation of errors between prediction and the target values, the computation of the gradient of the error function, and then the update of the weights. Seen also *feedforward neural network.*

Backward Chaining Backward chaining, also called goal-driven inference technique, is an inference approach that reasons backward from the goal to the conditions used to get the goal. Backward chaining inference is applied in many different fields, including game theory, automated theorem proving, and artificial intelligence.

Batch Normalization A preprocessing step where the data are centered around zero, and often the standard deviation is set to unity.

Bayesian Filter A Bayesian filter is a program using Bayesian logic. It is used to evaluate the header and content of email messages and determine whether or not it constitutes spam—unsolicited email or the electronic equivalent of hard copy bulk mail or junk mail. A Bayesian filter works with probabilities of specific words appearing in the header or content of an email. Certain words indicate a high probability that the email is spam, such as Viagra and refinance.

Bayesian Network A Bayesian Network, also called Bayes Network, belief network, or probabilistic directed acyclic graphical model, is a probabilistic graphical model (a statistical model) that represents a set of variables and their conditional dependencies via a directed acyclic graph (see *DAG*).

Biased algorithm See *Inadvertent effects of AI.*

Big Data The term big data is used when traditional data mining and handling techniques cannot uncover the insights and meaning of the underlying data. Data that are unstructured or time sensitive or simply very large cannot be processed by relational database engines. This type of data requires a different processing approach which uses massive parallelism on readily available hardware.

Blockchain Blockchain is a distributed system that records transactions across all users in an expanding chain of encrypted blocks. Blockchain builds a decentralized ledger that indicates every user has the same copy of the record. The records cannot be easily altered unless all of them are altered. Blockchain was invented in 2008 for the use of cryptocurrency bitcoin as a public transaction ledger. Such a system also shows its potential applications in different fields regarding the recording of events, medical records, and other record management systems.

Boolean neural network Boolean neural network is an artificial neural network approach which only consists of Boolean neurons (and, or, not). Such an approach reduces the use of memory space and computation time. It can be implemented to the programmable circuits such as FPGA (Field-Programmable Gate Array or Integrated circuit).

C

Caffe Caffe is short for Convolutional Architecture for Fast Feature Embedding which is an open source deep learning framework developed in Berkeley AI Research. It supports many different deep learning architectures and GPU-based acceleration computation kernels.

Case-Based Reasoning (CBR) Case-Based Reasoning is a way to solve a new problem by using solutions to similar problems. It has been formalized to a process consisting of case retrieve, solution reuse, solution revise, and case retention.

CE Marking A certification marking indicating conformity with standards for products sold within the European Economic Area. In the context of medical devices, CE Marking is similar to US Food and Drug Administration approval.

Central processing unit (CPU) Central processing unit is the electronic circuit within

that carries out the instructions of a computer program by performing the basic arithmetic, logical, control, and input/output operations specified by the instructions (see also *GPU*).

Chatbot Chatbot, also known as interactive agent, is an artificial intelligence system that uses natural language processing techniques to conduct a conversation via audio or texts. The most recognizable examples of chatbots are Apple's Siri, Microsoft's Cortana, and Amazon's Alexa.

Classification Classification is a general process for categorization which assigns a label to the samples. A classification system is an approach to accomplish categorization of samples.

Clinical Decision Support (CDS) A clinical decision support system is a health information technology system that is designed to provide physicians and other health professionals with clinical decision support, that is, assistance with clinical decision-making tasks.

Cloud The cloud is a general metaphor that is used to refer to the Internet. Initially, the Internet was seen as a distributed network and then with the invention of the World Wide Web as a tangle of interlinked media. As the Internet continued to grow in both size and the range of activities it encompassed, it came to be known as "the cloud." The use of the word cloud may be an attempt to capture both the size and nebulous nature of the Internet.

Cloud Computing Cloud Computing enables access to and usage of shared computer resources that can be provisioned with minimum management effort. The cloud is a general metaphor to refer to a group of networked computer resources that could provide computing services to avoid up-front IT infrastructures costs.

Clustering Clustering is a task to organize data into groups based on certain properties. Clustering analysis is widely used in data mining for pattern recognition, image analysis, and computer graphics, among others.

Cognitive computing Cognitive computing is used to refer to the systems that simulate the human brain to help with the decision-making. It uses self-learning algorithms that perform tasks such as natural language processing, image analysis, reasoning, and human–computer interaction. Examples of cognitive systems are IBM's Watson and Google DeepMind.

Cohort A sample in a clinical study (conducted to evaluate a machine learning algorithm, for example) where it is followed prospectively or retrospectively and subsequent status evaluations with respect to a disease or outcome are conducted to determine which initial participants' exposure characteristics (risk factors) are associated with it.

Computer-Aided Detection/Diagnosis (CAD) Computer-aided detection (CAD), or computer-aided diagnosis (CADx), uses computer programs to assist radiologists in the interpretation of medical images. CAD systems process digital images for typical appearances and highlight suspicious regions in order to support a decision taken by a professional.

Common Data Element (CDE) Common Data Element is a tool to support data management for clinical research.

Convolution The process of filtering. A filter (or equivalently: a kernel or a template) is shifted over an input image. The pixels of the output image are the summed product of the values in the filter pixels and the corresponding values in the underlying image.

Convolutional neural network (CNN) A convolutional neural network is a specific type of artificial neural network that uses *perceptrons*, a machine learning unit algorithm, for supervised learning, to analyze data. CNNs apply to image processing, natural language processing, and other kinds of cognitive tasks. A convolutional neural network is also known as a ConvNet. A CNN consists of an input and output layer as well as multiple hidden layers which are formed as mathematical operations. The hidden layers include convolutional layer, pooling layer, normalization, and fully connected layers. Since the success of AlexNet (see *Alexnet*) applied the ImageNet competi-

tion in 2013, there has been a rapid evolution of CNNs. VGGNet, GoogLeNet, ResNet, and DenseNet are some successful examples. See *Multilayer neural network*.

Computer Vision Computer Vision is an interdisciplinary field that uses computer science techniques to analyze and understand digital images and videos. Computer vision tasks include object recognition, event detection, motion detection, and object tracking, among others.

D

Data Data is a collection of qualitative and quantitative variables. It contains the information that is represented numerically and needs to be analyzed.

Data Cleaning Data Cleaning is the process of identifying, correcting, or removing inaccurate or corrupt data records.

Data Curation Data Curation includes the processes related to the organization and management of data which is collected from various sources.

Data-Driven Science Data-Driven Science, or Data Science, is an interdisciplinary field of employing computing algorithms to extract knowledge or insights from data acquired from different sources.

Data Extraction Data Extraction is the act or process of retrieving data out of data resources for further data processing or data storage.

Data Integration Data Integration involves the combination of data residing in different resources and then the supply in a unified view to the users. Data integration is in high demand for both commercial and scientific domains in which they need to merge the data and research results from different repositories.

Data Lake A type of data repository that stores data in its natural format and relies on various schemata and structure to index the data.

Data Mining Data Mining is the process of data analysis and information extraction from large amounts of datasets with machine learning, statistical approaches. and many others.

Deductive Reasoning Deductive Reasoning, also known as logical deduction, is a reasoning method that relies on premises to reach a logical conclusion. It works in a top-down manner, in which the final conclusion is obtained by reducing the general rules that hold the entire domain until only the conclusion is left.

Data Refinement Data refinement is used to convert an abstract data model in terms of sets for example into implementable data structures such as arrays.

Decision Tree A decision tree uses tree-like graph or model as a structure to perform decision analysis. It uses each node to represent a test on an attribute, each branch to represent the outcome of the test, and each leaf node to represent a class label.

Data Warehouse A data warehouse is typically an offline copy of production databases and copies of files in a non-production environment.

Deep Blue Deep Blue was a chess supercomputer developed by IBM. It was the first computer chess player that beat the world champion Garry Kasparov, after six-game match in 1997.

Deep Learning (DL) Deep Learning is a subfield of machine learning concerned with algorithms that are inspired by the human brain that works in a hierarchical way. Deep Learning models, which are mostly based on the (artificial) neural networks, have been applied to different fields, such as speech recognition, computer vision, and natural language processing.

DeepMind DeepMind is an artificial intelligence company founded in 2010 and later acquired by Google in 2014. DeepMind developed *Alphago* program that beat a human professional Go player for the first time.

Deep neural network A neural network architecture with many layers, typically 5–100. A network with only a few layers is called a shallow network.

Dice coefficient A measure to compare the similarity of two segmentations, e.g., by expert

and by machine. It is the ratio of twice the number of common pixels to the sum of all pixels in both sets.

Directed Acyclic Graph (DAG) In computer science and mathematics, a directed acyclic graph is a finite directed graph with no directed cycles. It consists of finitely many vertices and edges, with each edge directed from one vertex to another, such that there is no way to start at any vertex and follow a consistently directed sequence of edges that eventually loops back to that starting vertex again.

E

Electronic Medical Record (EMR) An electronic medical record, or electronic health record, is the systematized collection of patient and population electronically stored health information in a digital format. These records can be shared across different healthcare settings. Records are shared through network-connected, enterprise-wide information systems or other information networks and exchanges.

ELIZA The ELIZA effect is a term used to discuss progressive artificial intelligence. It is the idea that people may falsely attach meanings of symbols or words that they ascribe to artificial intelligence in technologies.

Enterprise Imaging Enterprise Imaging has been defined as "a set of strategies, initiatives and workflows implemented across a healthcare enterprise to consistently and optimally capture, index, manage, store, distribute, view, exchange, and analyze all clinical imaging and multimedia content to enhance the electronic health record" by members of the HIMSS-SIIM Enterprise Imaging Workgroup.

Error backpropagation The process of adjusting the weights in a neural network by minimizing the error at the output. It involves a large number of iteration cycles with the training data.

Ethics of Artificial Intelligence The ethics of artificial intelligence is the ethics of technology specific to robots and other artificial intelligence beings, which is divided into robot ethics and machine ethics. The former one is about the concern with the moral behavior of humans as they design, construct, use, and treat artificially intelligent beings, and the latter one is about the moral behavior of artificial moral agents (see also *inadvertent effects*).

Expert System Expert system is a computer system that simulates the ability or behavior of a human expert on performing a task. An expert system incorporates the knowledge base that represents facts and rules, and the inference engine that uses the knowledge base to deduce new conclusions.

Explainable artificial intelligence (XAI) Explainable artificial intelligence is a key term in AI design and in the tech community as a whole. It refers to efforts to make sure that artificial intelligence programs are transparent in their purposes and how they work. Explainable AI is a common goal and objective for engineers and others trying to move forward with artificial intelligence progress.

F

Fast Healthcare Interoperability Resources (FHIR) Fast Healthcare Interoperability Resources is a draft standard describing data formats and elements (known as "resources") and an application programming interface for exchanging electronic health records. The standard was created by the Health Level Seven International healthcare standards organization.

Forward Chaining Forward Chaining, also called forward reasoning, is a reasoning approach that searches inference rules from available data and then makes deduction and decision based on the rule. Forward Chaining works in the opposite as the backward chaining.

Feedforward Neural Network A feedforward neural network is an artificial neural network in which the connections between units do not form a cycle. The feedforward neural network has an input layer, hidden layers, and an output layer. Information always travels in one direction—from the input layer to the

output layer—and never goes backward. See also *backpropagation*.

Fully Convolutional Network (FCN) Fully Convolutional Network is the first convolutional neural network for semantic segmentation. It is trained end-to-end, pixel-to-pixel from arbitrary-sized inputs. Both learning and inference are performed whole image at a time by dense feedforward computation and backpropagation.

G

Generative Adversarial Network (GAN) A class or artificial intelligence algorithms used in unsupervised machine learning, where two neural networks (a generative network and a discriminative one) are pitted against one another—one network generates candidates, and the other evaluates them in a zero-sum game framework.

Genetic Algorithm Genetic algorithms are heuristic search and optimization algorithms inspired by the natural selection theory. A genetic algorithm requires a genetic representation of the solution and a fitness function to evaluate the solution.

Genomic data Genomic data refer to the genome and DNA data of an organism. They are used in bioinformatics for collecting, storing, and processing the genomes of living things. Genomic data generally require a large amount of storage and purpose-built software to analyze.

Gradient boost machine A type of machine learning technique that uses an ensemble of weak prediction models to perform regression and classification tasks.

Gradient descent A fast optimization method to find a minimum (e.g., error). The gradient is computed at the local position, and walking is done only a step in the downward direction. Repeating this process gives the fastest and most efficient way to the minimum.

Graphical Processing Unit (GPU) A graphical processing unit is a single chip processor designed for efficient manipulation of computer graphics and image processing, especially for computations that can be processed parallely. GPUs are widely used in embedded systems, mobile phones, personal computers, workstations, and many others. The rapid development of GPUs contributes to the rise of deep learning systems. The first GPU was developed by NVidia in 1999 and called the GeForce 256.

H

Heuristics A heuristic is a technique to provide fast or approximate solutions when the traditional methods are too slow or fail to give an accurate solution. A heuristic is commonly called a rule of thumb. While faster, it is typically less optimal than the classic methods it replaces.

Heuristic search techniques Support that narrows down the search for optimal solutions for a problem by eliminating options that are incorrect.

Human Level Machine Intelligence See: *Artificial General Intelligence.*

I

ImageNet ImageNet is a large image database with more than 14 million images over 20,000 categories. Since 2010, the ImageNet project runs annually a contest called the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) for object and scene recognition. The winner algorithm of the contest in the year 2012 is considered as the beginning of the deep learning revolution. See *AlexNet.*

Inadvertent effects of AI If training sets are poised with faulty data, then the algorithm will render faulty outcomes. A system is only as good as the data it learns from and databases must increase in order to let AI grow. See in the literature for racist, sexist algorithms. See also *Ethics of Artificial Intelligence.*

Inductive reasoning Inductive reasoning is a reasoning method which uses premises to supply evidence in order to support the conclusion. Opposed to deductive reasoning, inductive reasoning works as a down-top logic

which provides the conclusion by generalizing or extrapolating from special cases to general rules.

Interactive Machine Learning Interactive Machine Learning are approaches based on a coupling of human input and machines during the learning process.

Internet of Things (IoT) The Internet of Things (IoT) is the network of electronic devices embedded with softwares and sensors that enable the interaction between machines. The connectivity between devices helps the collection of huge data which can be analyzed by computer-based systems.

Internet of Medical Things (IoMT) Internet of Medical Things (IoMT) specifies the network of devices that are used to monitor the health status during the daily life.

Interoperability Interoperability is the property that allows for the unrestricted sharing of resources between different systems. This can refer to the ability to share data between different components or machines, or it can be defined as the exchange of information and resources between different computers through local area networks (LANs) or wide area networks (WANs). Broadly speaking, interoperability is the ability of two or more components or systems to exchange information and to use the information that has been exchanged.

Isaac Asimov Isaac Asimov (1920–1992) was a science fiction author and formulated the Three Laws of Robotics in the latter, which continues to influence researchers in robotics and artificial intelligence (AI).

K

Kaggle Kaggle is a data science platform to host data analysis competitions launched by companies and users.

Knowledge-Based Systems It is a computer system that uses knowledge to solve a problem or support a decision. A knowledge-based system has three types of subsystems: a knowledge base, a user interface, and an inference engine.

L

Label Also known as annotation. In supervised learning, the answer or result portion of an example. Each example in a labeled dataset consists of one or more features and a label. For instance, in a housing dataset, the features might include the number of bedrooms, the number of bathrooms, and the age of the house, while the label might be the house's price. In a spam detection dataset, the features might include the subject line, the sender, and the email message itself, while the label would probably be either spam or not spam.

Layer A layer, as in convolutional layer, is a set of neurons in a neural network that process a set of input features, or the output of those neurons. Deep learning networks get their name because they have many layers; most systems now have 30–150 layers, compared with traditional ANNs that would fail if they had more than about three layers.

Learning Learning is the process of acquiring new or modifying existing knowledge, behaviors, skills, values, or preferences. The ability to learn is possessed by humans, animals, and some machines, and there is also evidence for some kind of learning in some plants. Some learning is immediate, induced by a single event but much skill and knowledge accumulates from repeated experiences. See also *deep learning*, *machine learning, unsupervised* and *reinforcement learning.*

Learning algorithm A learning algorithm is an algorithm used in machine learning to help the technology to imitate the human learning process. Combined with technologies like neural networks, learning algorithms create involved, sophisticated learning programs.

Learning algorithm, examples Logic regression, linear regression, decision trees, and random forests are all examples of learning algorithms. Algorithms like "nearest neighbor" also involve the ways that these algorithms are used to affect decision-making and learning in machine learning. In general, what all of these algorithms have in common is their ability to extrapolate from test or training data to make projections or build models in the real

world. Think of these algorithms as tools for "pulling data points together" from a raw data mass or a relatively unlabeled background. Where learning algorithms are useful in both supervised and unsupervised machine learning, they are used in different ways in each type of discipline. Supervised machine learning benefits from having already labeled and isolated data, so the learning algorithms that are used will be different in some ways.

Learning rate A scalar used to train a model via gradient descent. During each iteration, the gradient descent algorithm multiplies the learning rate by the gradient. The resulting product is called the gradient step. Learning rate is a key hyperparameter.

Linear regression Linear regression is a kind of statistical analysis that attempts to show a relationship between two variables. Linear regression looks at various data points and plots a trend line. Linear regression can create a predictive model on apparently random data, showing trends in data. See *Learning algorithm, examples.*

Logistic regression Logistic regression is a kind of statistical analysis that is used to predict the outcome of a dependent variable based on prior observations. For example, an algorithm could determine the winner of a presidential election based on past election results and economic data. Logistic regression algorithms are popular in machine learning. See *Learning algorithm, examples.*

M

Machine intelligence See *Artificial Intelligence.*

Machine Learning Machine Learning is a field in computer science that builds computational models that have the ability of "learning" from the data and then provide predictions. Depending on whether there is a supervisory signal, machine learning can be divided into three categories: the *supervised learning*, *unsupervised learning*, and *reinforcement learning.*

Machine Vision Machine Vision is the technology used to provide image-based automatic analysis for applications in industry such as automatic inspection, process control, and robot guidance.

Markov Chain Any multivariate probability density whose independence diagram is a chain. In other words, the variables are ordered, and each variable "depends" only on its neighbors in the sense of being conditionally independent of the others. An equivalent definition is that you sample the variables left-to-right, conditional only on the last outcome.

Mask R-CNN Mask R-CNN is a general deep learning-based framework for object instance segmentation. It consists of two stages, in which the first stage performs a region proposal network that proposes candidate object bounding box, while the second stage provides a class prediction to the instances in the bounding box as well as a binary mask for instance segmentation.

Medical Imaging Informatics MII is the development, application, and assessment of information technology (IT) for clinical medical imaging. It includes the interfaces of IT and people. In practical terms, MII already occurs at a basic level throughout radiology practice, from the moment a clinician considers ordering an imaging study until images and interpretation are used to plan the patient's treatment.

Monte Carlo Methods Monte Carlo Methods, or Monte Carlo Simulation, are computational algorithms that rely on random sampling to obtain numerical results based on probability distributions. One example of using Monte Carlo Method is to approximate the value of π. It is done by uniformly scattering random points inside a square and then computing the ratio between the number of points falling in the circle and that of the total number of points within the square, which is equal to $\pi/4$.

Moore's Law Named after the cofounder of Intel, Moore predicted in 1965 that the number of transistors that can be placed on an integrated circuit doubles every 2 years. This trend has been continuing since 1965 with no signs of any slowdown yet. It can be

applied in general to a range of technology areas that are growing at an accelerating rate.

Multilayer neural network A multilayer neural network contains more than one layer of artificial neurons or nodes. They differ widely in design. It is important to note that while single-layer neural networks were useful early in the evolution of AI, the vast majority of networks used today have a multilayer model. Multilayer neural networks can be set up in numerous ways. Typically, they have at least one input layer, which sends weighted inputs to a series of hidden layers, and an output layer at the end. These more sophisticated setups are also associated with nonlinear builds using sigmoids and other functions to direct the firing or activation of artificial neurons. While some of these systems may be built physically, with physical materials, most are created with software functions that model neural activity. Convolutional neural networks (CNNs), used for image processing and computer vision, as well as recurrent neural networks, deep networks, and deep belief systems are all examples of multilayer neural networks. CNNs, for example, can have dozens of layers that work sequentially on an image. All of this is central to understanding how modern neural networks function.

Multilayer Perceptrons (MLP) A multilayer perceptron is a class of feedforward artificial neural network. An MLP consists of at least three layers of nodes. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. MLP utilizes a supervised backpropagation technique for training. Its multiple layers and nonlinear activation distinguish MLP from a linear perceptron. It can distinguish data that is not linearly separable.

Multi-task Learning Multi-task learning (MTL) is a subfield of machine learning in which multiple learning tasks are solved at the same time while exploiting commonalities and differences across tasks. This can result in improved learning efficiency and prediction accuracy for the task-specific models, when compared to training the models separately.

N

Narrow artificial intelligence (narrow AI) Narrow artificial intelligence (narrow AI) is a specific type of artificial intelligence in which a technology outperforms humans in some very narrowly defined task. Unlike general artificial intelligence, narrow artificial intelligence focuses on a single subset of cognitive abilities and advances in that spectrum.

Natural Language Processing Natural language processing (NLP) is a method to translate between computer and human languages. Traditionally, feeding statistics and models have been the method of choice for interpreting phrases. Recent advances in this area include voice recognition software, human language translation, information retrieval, and artificial intelligence. There is difficulty in developing human language translation software because language is constantly changing. Natural language processing is also being developed to create human readable text and to translate between one human language and another. Already existing reports associated with radiology images can be used to learn about disease and conditions and the ultimate goal of NLP is to build software that will analyze, understand, and generate human languages naturally, enabling communication with a computer as if it were a human.

Neural networks Also known as artificial neural network, neural net, deep neural net; a computer system inspired by living brains. Neural networks found to perform best in ImageNet data challenges were *convolutional* neural networks (CNNs). This name comes from the mathematical concept of convolution, which is similar to the CNN convolutional operation wherein filters are applied to an image in fixed spatial regions and are swept across, or integrated, over the entire image. The resulting activations can then be aggregated in pooling operations,

subjected to repeated convolutions, and eventually mapped to a vector of probabilities corresponding to likelihoods that the image belongs to a certain class.

O

Omics The word omics indicates the study of a body of information and refers to the fields of biology ending in -omics such as genome, proteome, microbiome, and exposome. Many of the emerging fields of large-scale data-rich biology are designated by adding the suffix -omics onto previously used terms.

OpenAI OpenAI is a nonprofit artificial intelligence research company (founded in December 2015 by partners including Elon Musk) that aims to promote and develop friendly AI in such a way as to benefit humanity as a whole. The organization aims to "freely collaborate" with other institutions and researchers by making its patents and research open to the public.

Overfitting In statistics and machine learning, overfitting occurs when a model tries to predict a trend in data that is too noisy. Overfitting is the result of an overly complex model with too many parameters. A model that is overfitted is inaccurate because the trend does not reflect the reality of the data. An overfitted model is a model with a trend line that reflects the errors in the data that it is trained with, instead of accurately predicting unseen data. This is better seen visually with a graph of data points and a trend line. An overfitted model shows a curve with higher and lower points, while a properly fitted model shows a smooth curve or a linear regression.

Overfitting, compensation of Overfitting typically results from an excessive number of training points. There are a number of techniques that machine learning researchers can use to mitigate overfitting, including cross-validation, regularization, early stopping, pruning, Bayesian priors, dropout, and model comparison.

P

Pattern matching Pattern recognition and pattern matching are sometimes confused as the same thing when, in fact, they are not. Whereas pattern recognition looks for a similar or most likely pattern in a given data, pattern matching looks for exactly the same pattern. Pattern matching is not considered part of *machine learning*, although in some cases it leads to similar results as pattern recognition. Pattern recognition has its origins in engineering, whereas machine learning grew out of computer science. Both can be viewed as two facets of the same field.

Pattern recognition In IT pattern recognition is a branch of *machine learning* that emphasizes the recognition of data patterns or data regularities in a given scenario. It is a subdivision of *machine learning* and it should not be confused with actual machine learning study. Pattern recognition can be either supervised, where previously known patterns can be found in a given data, or unsupervised, where entirely new patterns are discovered. The objective behind pattern recognition algorithms is to provide a reasonable answer for all possible data and to classify input data into objects or classes based on certain features. A most likely matching is performed between various data samples and their key features are matched and recognized.

Perceptron The perceptron computes a single output from multiple real-valued inputs by forming a linear combination according to its input weights and then possibly putting the output through some nonlinear activation function. A multilayered perceptron is a network of simple neurons called perceptrons. The basic concept of a single perceptron was introduced by Rosenblatt in 1958.

Perceptron algorithm Perceptron algorithm is a machine learning algorithm that helps provide classified outcomes for computing. Perceptron algorithm is called supervised classification because the computer is aided by the human classification of data points. Perceptron is also related to the development of "artificial neural networks,"

where computing structures are based on the design of the human brain.

Planning A branch of AI dealing with planned sequences or strategies to be performed by an AI-powered machine. Things such as actions to take, variable to account for, and duration of performance are accounted for.

Principal component analysis (PCA) Constructing new features which are the principal components of a dataset. The principal components are random variables of maximal variance constructed from linear combinations of the input features. Equivalently, they are the projections onto the principal component axes, which are lines that minimize the average squared distance to each point in the dataset. To ensure uniqueness, all of the principal component axes must be orthogonal. PCA is a maximum-likelihood technique for linear regression in the presence of Gaussian noise on both inputs and outputs. In some cases, PCA corresponds to a Fourier transform, such as the DCT used in JPEG image compression.

Pruning The use of a search algorithm to cut off undesirable solutions to a problem in an AI system. It reduces the number of decisions that can be made by the AI system.

Python Programming language that runs on most platforms and is often used for data science, machine learning, and deep learning.

R

Radiomics The -omics of images is an expansion of CADx. Radiomics refers to the extraction and analysis of large amounts of advanced quantitative image features with the intent of creating mineable databases from radiological images. From which prognostic associations can be made between images and outcomes.

Radiogenomics This term is used in two contexts. Either to refer to the study of genetic variation associated with response to radiation or to refer to the correlation between cancer imaging features and gene expression. It is the combination of radiomics and genomics, the gene profile of, for example, a tumor. Combining both radiomics and radiogenomics will lead to AI predicting which kind of gene profile defect there is based on its features seen on scans.

Random Forests (or Random Decision Forests) Random Forests or Random Decision Forests are ensembling learning methods for data classification and regression. They construct a multitude of *decision trees* during the training and output the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees.

Receptive field (RF) The sensitivity pattern of a neuron. For example, the receptive field of a simple cell in the primary visual cortex V1 is determined by measuring its firing rate as a function of the pointwise scanning light stimulation of its receptive field area on the retina. A receptive field is the biological implementation of a filter.

Recurrent neural network (RNN) A type of neural network that makes sense of sequential information and recognizes patterns, and creates outputs based on those calculations. Remembers a previous state in its memory, and feeds this back as one of the inputs. It is characterized by a recurrent loop in the architecture. This type of neural network is used for sequential data, e.g., text and video.

Regression Regression is a process of predicting the value to a yes or no label provided it falls on a continuous spectrum of input values, subcategory of *supervised learning*.

Reinforcement Learning Reinforcement learning is a type of dynamic programming that trains algorithms using a system of reward and punishment. The algorithm is exposed to a total random and new dataset and it automatically finds patterns and relationships inside of that dataset. The system is rewarded when it finds a desired relationship inside of that dataset but it is also punished when finds an undesired relation. The algorithm learns from awards and punishments and updates itself continuously. This type of algorithm is

always in production mode. It requires real-time data to be able to update and present actions. The agent learns without intervention from a human by maximizing its reward and minimizing its penalty.

Residual neural network (RNN) This network skips connections over network layers, by making shortcuts or jump-overs. A ResNet skips over a single layer.

S

Scikit-learn Scikit-learn (formerly scikits-learn) is a free software machine learning library for the *Python* programming language. It features various classification, regression, and clustering algorithms, including support vector machines, random forests, gradient boosting, k-means, and DBSCAN, and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy.

Singularity Singularity refers to the emergence of superintelligent machines with capabilities that cannot be predicted by humans. The word singularity comes from astrophysics where it is used to refer to a point in space time where the rules of ordinary physics do not apply. This idea is parallel to the way the term is used in a technological context, because if a technological singularity were to occur, humans would be become unable to predict events beyond that point. See *Superintelligence.*

Strong AI An area of AI development that is working toward the goal of making AI systems that are as useful and skilled as the human mind.

Stride The step size in the shift of convolution filters. It is normally set to 1, but can be 2–10 or even higher, to increase the computational efficiency.

Supervised Learning Training a model from input data and its corresponding labels. Supervised machine learning is analogous to a student learning a subject by studying a set of questions and their corresponding answers. After mastering the mapping between questions and answers, the student can then provide answers to new questions on the same topic. See also *unsupervised machine learning*.

Support Vector Machine (SVM) Support Vector Machine, or in short SVM, is a supervised machine learning model for data classification and regression analysis. One of the most used classifiers in machine learning. It optimizes the width of the gap between the points of separate categories in feature space.

Superintelligence A superintelligence is an intelligence system that rapidly increases its intelligence in a short time, specifically, to surpass the cognitive capability of the average human being. Part of the idea of superintelligence is that certain kinds of artificial intelligence work are theoretically capable of triggering a "runaway reaction" where an artificial intelligence far exceeds human capacity for thought and starts to manipulate or control humans in specific ways. Superintelligence is tied to the idea of a "singularity," which is based on the idea that a catalyst or trigger would cause rapid change beyond what humans can anticipate. See *Singularity.*

T

TensorFlow TensorFlow is an open source software library for numerical computation using data flow graphs. Nodes in the graph represent mathematical operations, while the graph edges represent the multidimensional data arrays (tensors) communicated between them. The flexible architecture allows you to deploy computation to one or more CPUs or GPUs in a desktop, server, or mobile device with a single API. TensorFlow was originally developed by researchers and engineers working on the Google Brain Team within Google's Machine Intelligence research organization for the purposes of conducting machine learning and deep neural networks research, but the system is general enough to be applicable in a wide variety of other domains as well.

Tensors Multidimensional arrays of primitive data values that are used in TensorFlow. A

tensor consists of a set of primitive values shaped into an array of any number of dimensions. These massive numbers of large arrays are the reason that GPUs and other processors designed to do floating point mathematics excel at speeding up these algorithms.

Tensor Processing Unit A unit similar to a Graphic Processing Unit, it is a measure of tensor processing power.

Turing test A test developed by Alan Turing in the 1950s that tests the ability of a machine to mimic human behavior (see terms "*Computing Machinery and Intelligence*"). The test involves a human evaluator who undertakes natural language conversations with another human and a machine and rates the conversations. It is designed to determine whether or not a computer could be classed as intelligent. The test (also referred to as the imitation game) is conducted by having human judges chat to several people via a computer. Most of the people the judges will be speaking to are humans, but one will actually be a chatbot. The chatbot's objective will be to convince the human judges that they are speaking to a real person. If it does this, it has passed the Turing test.

U

Uncanny valley The uncanny valley is a phenomenon that occurs in the human psyche and perception with regard to objects that are human-like, usually robots and images, and determines our reaction toward that object. It is still just a hypothesis, and it is stated to the effect of "as an object such as a robot gets more human-like, the response of some observers will become increasingly positive and emphatic, until a point is reached in the robot's human-likeness beyond which the reactions turn to strong revulsion."

U-net A network with a U-shape, where connections exist between the horizontally corresponding layers of the contracting input branch and the expanding output branch. It was designed to work with fewer training images and to yield more precise segmentations.

Unsupervised learning Unsupervised learning is a type of machine learning algorithm used to draw inferences from sets of data consisting of input data without labeled responses, e.g., cluster analysis. This means that the system is exposed to a total random and new dataset and it automatically finds patterns and relationships inside of that dataset. Unsupervised learning is used in email clustering in order to distinguish between spam emails and useful emails. It can also be seen as Learning by Example. Another example of unsupervised machine learning is *principal component analysis* (PCA). For example, applying PCA on a dataset containing the contents of millions of shopping carts might reveal that shopping carts containing lemons frequently also contain antacids.

Underfitting Underfitting occurs when a statistical model cannot adequately capture the underlying structure of the data.

V

Variational Autoencoder Variational autoencoder (VAE) models inherit autoencoder architecture, but make strong assumptions concerning the distribution of latent variables. They use variational approach for latent representation learning, which results in an additional loss component and specific training algorithm called Stochastic Gradient Variational Bayes. It assumes that the data is generated by a directed graphical model and that the encoder is learning an approximation to the posterior distribution and denote the parameters of the encoder (recognition model) and decoder (generative model), respectively.

W

Watson Watson is named after Dr. Watson, a former IBM CEO. It is a question-answering supercomputer that uses AI to perform cognitive computing and data analysis. In the year 2011, Watson competed on the *Jeopardy!*

television show against human contestants and won the first place prize. Since then, Watson has been used for utilization management in medical centers.

Weak AI See: *Artificial Narrow Intelligence.*

Weights The connection strength (coefficients) between units or nodes in a neural network. These weights can be adjusted in a process called *learning*. The goal of training a linear model is to determine the ideal weight for each feature. If a weight is 0, then its corresponding feature does not contribute to the model.

Winters See *Artificial winters.*

Index

E. R. Ranschaert et al. (eds.), *Artificial Intelligence in Medical Imaging*,
https://doi.org/10.1007/978-3-319-94878-2