IMDRF
International Medical Device Regulators Forum

IMDRF/AIMD WG/N67

Machine Learning-enabled Medical Devices: Key Terms and Definitions

AUTHORING GROUP

Artificial Intelligence Medical Devices (AIMD) Working Group

6 May 2022

Preface

Tracey Duffy, IMDRF Chair

Contents

1. Introduction

Artificial Intelligence (AI) is a branch of computer science, statistics, and engineering that uses algorithms or models to perform tasks and exhibit behaviors such as learning, making decisions and making predictions. The subset of AI known as Machine Learning (ML) allows ML models to be developed by ML training algorithms through analysis of data, without models being explicitly programmed.

Approaches utilizing ML, sometimes colloquially referred to as AI or AI/ML, have been employed in several fields, such as the automotive industry, robotics, medicine, finance, and art. ML has given many sectors an ability to gain new insights from large amounts of data and to support tasks.

Examples in healthcare applications include earlier disease detection and diagnosis; identification of new observations or patterns on human physiology; development of personalized diagnostics and therapeutics; workflow optimization; signal processing and reconstruction; and guidance in use of the device with the goal of improving user and patient experience. There has been accelerated adoption and use of ML-enabled approaches in medical devices. We refer to these medical devices as Machine Learning-enabled Medical Devices, or MLMD. AI-based systems are typically implemented as software in medical devices or as Software as a Medical Device. MLMD have the potential to transform health care by deriving new and important insights from the vast amount of data generated during all phases of the healthcare process. One of the greatest benefits of MLMD resides in the opportunity for further learning and iteration as additional data becomes available, including from real-world use and experience to improve its performance.

The purpose of this publication is to establish relevant terms and definitions across the Total Product Life Cycle (TPLC) to promote consistency, support global harmonization efforts, and provide a foundation for the development of future guidelines related to MLMD. Terms referenced herein have either been previously defined in Global Harmonization Task Force (GHTF) documents or by internationally recognized standards on AI. Some terms and definitions have been generated by or are discussed by the IMDRF Artificial Intelligence Medical Devices (AIMD) Working Group within this document.

The overarching objective of this effort is to promote consistent expectations and understanding for MLMD, promote patient safety, foster innovation, and encourage access to advances in healthcare technology.

2. Scope

This document applies to key terms and definitions relating to Machine Learning-enabled Medical Devices (MLMD).

Note 1 : MLMD are *medical devices*. A product must first meet the definition of a *medical device* before it can be an MLMD.

Note 2 : Most jurisdictions include "accessories to medical devices" in the definition of "medical device". Other jurisdictions define "accessories to medical devices" separately. The definitions and the concepts in this document are intended to apply in both cases.

Note 3 : This document does not attempt to define established terms in the field of computer science; however, it does strive to highlight and clarify conflicting terms and definitions as necessary. This document does not provide guidelines for the development, risk management or evaluation of MLMD.

Note 4 : Terms and definitions that refer to technical standards that are under development (e.g. ISO, IEC, IEEE) may be updated upon final publication of those standards.

3. References

3.1. IMDRF / GHTF

- IMDRF/SaMD WG/N10 FINAL:2013 Software as a Medical Device (SaMD): Key Definitions
- IMDRF/GRRP WG/N47:2018 Essential Principles of Safety and Performance of Medical Devices and IVD Medical Devices(3.0 Definitions)

3.2. Standards

The standards below were consulted in the writing of this document and may be useful in meeting the key definition of MLMD discussed herein. This list is not intended as a required or complete list of standards that can be used to meet the key definition of MLMD.

- ISO/IEC DIS 22989 Information technology — Artificial intelligence — Artificial Intelligence Concepts and Terminology
- ISO/IEC TR 24027 Information technology — Artificial intelligence (AI) — Bias in AI systems and AI aided decision making

3.3. Other Documents

- AAMI, BSI, Turpin, R., Hoefer, E., Lewelling, J., & Baird, P. (2020). *Machine Learning AI in Medical Devices: Adapting Regulatory Frameworks and Standards to Ensure Safety and Performance.* AAMI/BSI Initiative on Artificial Intelligence.
  https://www.bsigroup.com/en-US/medical-devices/resources/ Whitepapers-and-articles/machine-learning-ai-in-medical-devices/

- Kohavi, R., & Provost, F. (Eds.). (n.d.). Glossary of Terms: Special Issue on Applications of Machine Learning and the Knowledge Discovery Process.
  https://ai.stanford.edu/~ronnyk/glossary.html

- Kan A. (2017). Machine learning applications in cell image analysis. Immunology and Cell Biology, 95(6), 525–530.
  https://doi.org/10.1038/icb.2017.16

4. General Overview of Artificial Intelligence and Machine Learning Concepts

AI-based systems are able to perform tasks such as visual perception, speech recognition, decision-making, and translation between languages – by using expert systems (based on rules like decision trees), machine learning (for example deep learning).

Some AI-based systems demonstrate a degree of autonomy (level of capacity to perform tasks in a complex environment without constant guidance/input from a user) and a capacity for adaptability (extent of the ability to learn from experience and thereby change performance).

ML involves a computer implementing an ML training algorithm to learn patterns from data, including classification, inference, matching previous patterns, predicting future outputs, etc., which results in an ML model to be applied to new data. ML has been considered as a subset of AI that gives computers the ability to learn without being explicitly programmed.
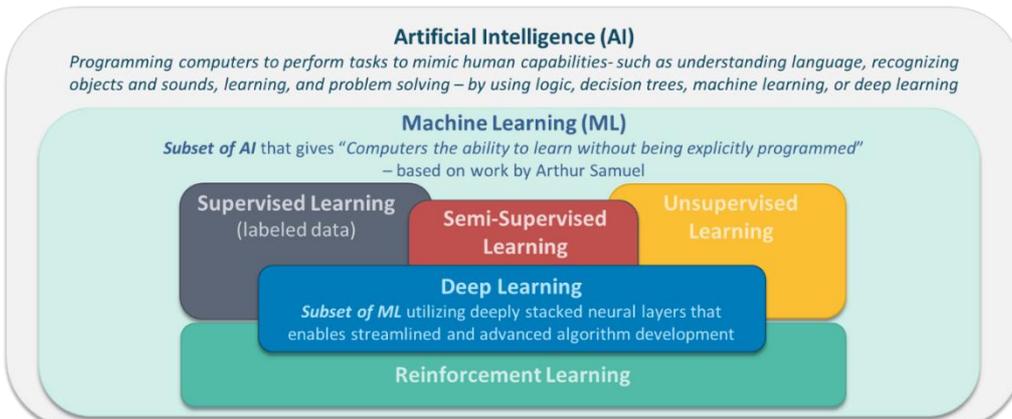
※ The descriptions within the diagram are not definitions, and are included to convey a general sense of the technology.

Figure 1 Overview of AI and ML Concepts

ISO/IEC's draft international standard for AI, DIS 22989, describes ML as a process that uses computational techniques to optimise model operation such that the ML model's behaviour reflects the data or experience.

There are several different types of ML methods (Figure 1), as well as different ML training algorithms. For example, some applications may use Supervised Learning, others may use Unsupervised or Semi-Supervised Learning (Section 6.0). Still others may use an iterative process of trial and error, also known as reinforcement learning[1]. Different types of ML training algorithms include neural networks (e.g. feedforward neural network, recurrent neural network, convolutional neural network, etc.) bayesian networks, decision trees, and support vector machines, among others.

Note : Within this document, the term ML training algorithm is used to represent a software procedure that establishes the parameters of a machine learning model by analyzing data. The term ML model is used in this document to refer to a mathematical construct that generates an inference or prediction based on new input data, and is the result of an ML training algorithm learning from data.

The following sections provide key definitions that are relevant to ML when used in medical devices (Section 5.0) and definitions from technical standards (Section 6.0), followed by a discussion of common ML terms (Section 7.0).

---

[1] Reinforcement learning (RL) is learning by interacting with an environment. An reinforcement learning model learns from the consequences of its actions, rather than from being explicitly taught and it selects its actions on basis of its past experiences (exploitation) and also by new choices (exploration), which is essentially trial and error learning. (Modified from 'http://www.scholarpedia.org/article/Reinforcement_learning'.)

5. Key Definitions

5.1. Machine Learning-enabled Medical Device (MLMD)

A medical device that uses machine learning, in part or in whole, to achieve its intended medical purpose.

5.2. IMDRF Terms

Medical Device: Any instrument, apparatus, implement, machine, appliance, implant, reagent for in vitro use, software, material or other similar or related article, intended by the manufacturer to be used, alone or in combination, for human beings, for one or more of the specific medical purpose(s) of:

- diagnosis, prevention, monitoring, treatment or alleviation of disease,
- diagnosis, monitoring, treatment, alleviation of, or compensation for, an injury,
- investigation, replacement, modification, or support of the anatomy, or of a physiological process,
- supporting or sustaining life,
- control of conception,
- cleaning, disinfection or sterilization of medical devices,
- providing information by means of in vitro examination of specimens derived from the human body;

and does not achieve its primary intended action by pharmacological, immunological, or metabolic means, in or on the human body, but which may be assisted in its intended function by such means.

> **Note 1 :** Products which may be considered to be medical devices in some jurisdictions but not in others include:
>
> - disinfection substances,
> - aids for persons with disabilities,
> - devices incorporating animal and/or human tissues,
> - devices for in-vitro fertilization or assisted reproduction technologies.
>
> **Note 2 :** For clarification purposes, in certain regulatory jurisdictions, devices for cosmetic/aesthetic purposes are also considered medical devices.
>
> **Note 3 :** For clarification purposes, in certain regulatory jurisdictions, the commerce of devices incorporating human tissues is not allowed.

Editorial issue has been corrected from IMDRF/GRRP WG/N47:2018.

6. Definitions/Reference Definitions/Technical Standards Definitions

6.1. Bias

Systematic difference in treatment[2] of certain objects, people, or groups in comparison to others.

Note 1 to entry: Treatment is any kind of action, including perception, observation, representation, prediction or decision. (ISO/IEC TR 24027:2021)

Note : The term 'Bias' is used in different ways in different fields. For example, in data science, bias is often defined with a statistical/mathematical meaning while in law, bias is often used to mean unfair or unfairly prejudiced/partial.

The ISO/IEC TR 24027 definition is a technical definition and is not synonymous with notions of being 'unfair' or not. Further information on the differences between bias and fairness is available in ISO/IEC TR 24027:2021.

ISO/IEC TR 24027 refers to systems having both "wanted" and "unwanted" bias depending on the intended purpose of an AI(-based) system. For instance, for an MLMD intended for the detection of leukemia, a wanted bias, would be bias toward the detection of leukemia over other pathologies; unwanted bias may include unintended differences in performance across different age groups in the intended patient population. As such, and depending on intended purpose, an MLMD that is more effective at the detection of leukemia in one age group over another might be an example of a device that has "unwanted" bias.

Sources of bias include:

- human cognitive biases (including automation bias, societal bias, and confirmation bias),
- data biases (including statistical bias, data processing bias, and data aggregation bias), and
- bias introduced by engineering decisions (e.g., during feature engineering, via algorithm selection, and model bias)

Further information on the types, and sources, of bias is provided in ISO/IEC TR 24027.

---

[2] "Treatment" in this definition does not infer or limit to medical or clinical treatment, the term more broadly refers to any kind of action, including perception, observation, representation, prediction or decision (ISO/IEC TR 24027:2021)

6.2. Continuous Learning

Training that leads to change of an MLMD with each exposure to data that takes place on an ongoing basis during the operation phase of the MLMD life cycle. (Modified from ISO/IEC DIS 22989)

Note : Although not necessarily in opposition, Batch Learning is often referred to when describing Continuous Learning. Batch Learning is a training that leads to the change of an MLMD that involves discrete updates based on defined sets of data that take place at distinct points prior to or during the operation phase of the MLMD life cycle.

6.3. Reference Standard

An objectively determined benchmark that is used as the expected result for comparison, assessment, training, etc.

6.4. Reliability

Property of consistent intended behavior and results. (ISO/IEC DIS 22989)

6.5. Semi-Supervised Machine Learning

Machine learning algorithms that leverage both unsupervised and supervised techniques during training. (Modified from ISO/IEC DIS 22989)

Note 1 : Descriptive information can be broader than just labelling. Annotation is the process of attaching descriptive information to data, such as metadata, labels, or anchors. The data itself is unchanged in the annotation process[3].

Note 2 : Additional information about this term can be found in Section 7.2

6.6. Supervised Machine Learning

Machine learning that makes use of labelled data during training. (ISO/IEC DIS 22989)

Note 1 : Descriptive information can be broader than just labelling. Annotation is the process of attaching descriptive information to data, such as metadata, labels, or anchors. The data itself is unchanged in the annotation process[3].

Note 2 : Additional information about this term can be found in Section 7.2

6.7. Test Dataset

 A set of data that is never shown to the ML training algorithm during training, that is used to estimate the ML model's performance after training.

6.8. Training

Process intended to establish or to improve the parameters of a ML model, based on an ML training algorithm, by using training data. (Modified from ISO/IEC DIS 22989)

---

[3] ISO/IEC DIS 22989 Information technology — Artificial intelligence — Artificial Intelligence Concepts and Terminology

IMDRF International Medical Device Regulators Forum

6.9. Training Dataset

 A set of data that is used to train the ML model, which is not part of the Test Dataset.

6.10. Unsupervised Machine Learning

Machine learning that only makes use of unlabelled data during training. (Modified from ISO/IEC DIS 22989)

Note 1 : Additional information about this term can be found in Section 7.2

7. Discussion

The following sub-sections contain discussions of concepts that warranted more detail than a concise definition. In particular, the aspects of MLMD changes, supervised and unsupervised learning, and validation are discussed.

7.1. Aspects of MLMD Changes

MLMD offer unique benefits, flexibility, and challenges related to their capacity for change. The transparent communication of the various aspects of these changes is important to the safety, performance, and effectiveness of MLMD.

The examples outlined in this discussion are not exhaustive and the relevant information may expand over time. It is important to note that changes, such as software patches, operating system updates, cybersecurity improvements, etc., can impact both MLMD and non-MLMD and, although important, these changes are not within the scope of this discussion.

There are a number of unique changes related to MLMD, including changes to the ML model or to the environment of use relative to the ML training data. The following discussion highlights these important aspects in two sections, MLMD Changes and MLMD Environmental Changes.

7.1.1. Changes to MLMD

Aspects that describe changes to MLMD include the cause, effect, trigger, domain, and effectuation. These attributes describe what changes, as well as why, where, when, and how the MLMD change occurs. MLMD is in a locked state when changes are not permitted.

Note : The word "locked" has been used by the community in a number of different ways. Some have defined a "locked device" as one that has been developed using ML methods but for which the developer does not have an intention of modifying at the present time. Others have used the term "locked device" as any device that does not perform "continuous learning." When using the word "locked" it is important to provide clarifying language around its use to communicate how it is being used.
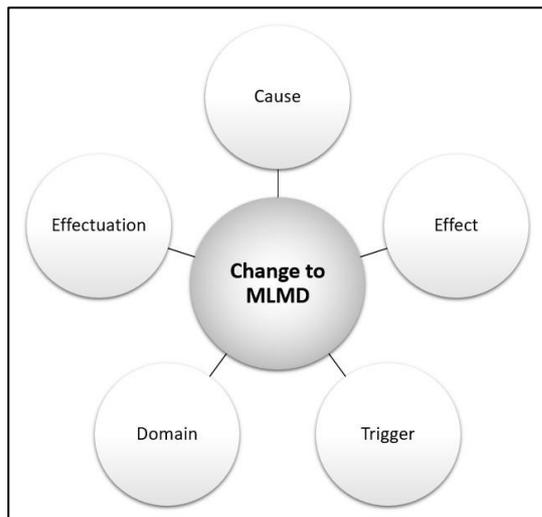

Figure 2 Aspects of MLMD Changes

The **cause** refers to the source of the change to the MLMD, for example, re-training with new or appended data different training methods or ML training algorithms, additional ML model, tuning, etc.

The **effect** refers to the resulting change to the MLMD, which can include amended intended use/indications for use; modified performance, changes in inputs, outputs, etc.

The **trigger** refers to the event that prompts or instigates the change to the MLMD, which can include performance thresholds, training data batch-size thresholds, exposure to new data/experiences, scheduled time intervals, MLMD environmental changes, user feedback, etc.

The **domain** refers to the scope or applicable extent of the change to the MLMD, which can be categorized as either homogeneous or heterogeneous. A homogeneous change is a uniform change that occurs universally (sometimes referred to as a global adaptation, note that global does not denote around-the-world). Heterogeneous changes are non-uniform changes that can be specific to one clinic, region, demographic, etc. (sometimes referred to as local adaptations)[4].

The **effectuation** refers to where the mechanism for change implementation resides, which can either be external (i.e., updated by the developer or user) or internal (i.e., updated by a change-control-software within the device).

---

[4] "Introduction to Online Machine Learning: Simplified",
https://www.analyticsvidhya.com/blog/2015/01/introduction-online-machine-learning-simplified-2/

7.1.2. Changes to MLMD Environment for Data

An MLMD environmental change is a modification to the setting of the MLMD relative to the ML development data. Aspects that describe an MLMD environmental change include the cause, effect, and domain.
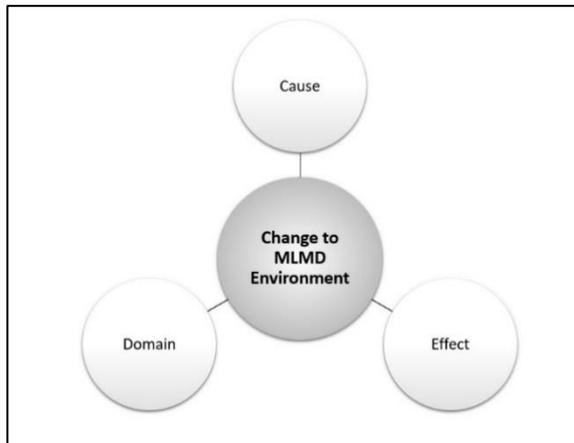

Figure 3 Aspects of MLMD Environmental Changes

The **cause** of an MLMD environmental change refers to the source of the change relative to the development environment. Examples of such causes include changes to the format or quality of the MLMD inputs (e.g., changes to third party image processing, incidents of adversarial machine learning); changes in the patient population (e.g., demographic shift); changes in clinical practice (e.g., earlier interventions that mask features used by the ML model for classification), etc.

The **effect** of an MLMD environmental change can involve deteriorated or improved performance, effectiveness, or safety.

The **domain** of an MLMD environmental change refers to the scope or applicable extent of the change, which can be categorized as either homogeneous or heterogeneous. Heterogeneous changes are non-uniform changes that can be specific to one clinic, region, demographic, etc. (sometimes referred to as local changes). Homogeneous changes are changes that occur uniformly (universally, globally) over some groups or settings/context. Note that global does not denote around-the-world.

7.2. Supervised / Unsupervised / Semi-Supervised Learning

Supervised and Unsupervised Machine Learning are two methods that are commonly used to ML models, but they are not the only methods available. The terms "supervised" and "unsupervised" in a machine learning context refer to the training methods, and specifically whether labelled or unlabelled data are used. Supervised Machine Learning utilizes labelled data during Training to learn the relationship between independent attributes and a designated dependent attribute (the label). In other words, supervised learning is a task to learn a mapping from input to output values, where the correct output values are known (labelled training data). Examples of supervised learning include decision trees, bayesian models, and regression analyses. Unsupervised Machine Learning utilizes unlabelled data during Training to group data without a pre-specified dependent attribute. In other words, unsupervised learning is the ability to find patterns from input values, where the output values are unknown. Examples of unsupervised learning include some types of ML training algorithms that perform clustering or dimensionality reduction.

Machine learning systems can use a mix of supervised and unsupervised learning (sometimes referred to as semi-supervised learning), as well as other learning methods such as reinforcement learning.

The terms "Supervised Machine Learning" and "Unsupervised Machine Learning" are often misunderstood. When used in a machine learning context, "supervised" or "unsupervised" does not refer to the presence or absence of a human supervisor overseeing the software. "Supervised" or "unsupervised" does not refer to the role that the software plays in a clinical environment, i.e., it does not describe the level of "autonomy" in practice. "Supervised" or "unsupervised" also does not refer to whether the software updates itself in a self-effectuating update process, i.e., whether it performs its own updates or adaptations.

7.3. Validation

The term validation has been used to represent different concepts within the fields of medical device development and ML model development.

Validation within the context of medical device development has been defined as follows:

> *Validation means confirmation by examination and provision of objective evidence that the particular requirements for a specific intended use can be consistently fulfilled[5].*

The term validation has also been used within the field of machine learning to refer to either data curation (sometimes referred to as data validation) or ML model tuning[6].

Data curation and ML model tuning can occur throughout the product lifecycle. Data curation refers to the selection, management and assessment of the independent and dependent attributes (labels) of data sets. ML model tuning is a particular phase of model development during which ML model is tuned; this optional tuning phase can be combined with the Training phase to optimize the ML model selection.

MLMD manufacturers, regulators, and users should be aware of the conflicting interpretations of the term validation and ensure that communication regarding the development phases and the associated datasets is clear to avoid confusion between data validation, ML model tuning, and medical device validation. Alternatively, the use of the term validation that refers to the training and tuning process should be avoided in the context of medical device development. It is recommended that the use of the term "validation" be accompanied by the context when referring to ML model tuning, data curation, and the associated datasets.

---

[5] Design Control Guidance for Medical Device Manufacturers (GHTF.SG3.N99-9)

[6] Ripley, B. (1996). Glossary. In Pattern Recognition and Neural Networks (pp. 347-354). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511812651.013

IMDRF International Medical Device
Regulators Forum

Please visit our website
for more details.

www.imdrf.org