INTERNATIONAL STANDARD

ISO/IEC 23053

First edition
2022-06

Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)

Cadre méthodologique pour les systèmes d'intelligence artificielle (IA) utilisant l'apprentissage machine

Reference number
ISO/IEC 23053:2022(E)

© ISO/IEC 2022



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2022

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

© ISO/IEC 2022 – All rights reserved

Contents

Page

Foreword ... **iv**

Introduction ... **v**

1 Scope ... **1**

2 Normative references ... **1**

3 Terms and definitions ... **1**
3.1 Model development and use ... 1
3.2 Tools ... 2
3.3 Data ... 2

4 Abbreviated terms ... **3**

5 Overview ... **4**

6 Machine learning system ... **4**
6.1 Overview ... 4
6.2 Task ... 5
6.2.1 General ... 5
6.2.2 Regression ... 6
6.2.3 Classification ... 6
6.2.4 Clustering ... 6
6.2.5 Anomaly detection ... 6
6.2.6 Dimensionality reduction ... 7
6.2.7 Other tasks ... 7
6.3 Model ... 7
6.4 Data ... 8
6.5 Tools ... 9
6.5.1 General ... 9
6.5.2 Data preparation ... 9
6.5.3 Categories of ML algorithms ... 10
6.5.4 ML optimisation methods ... 14
6.5.5 ML evaluation metrics ... 16

7 Machine learning approaches ... **19**
7.1 General ... 19
7.2 Supervised machine learning ... 20
7.3 Unsupervised machine learning ... 22
7.4 Semi-supervised machine learning ... 23
7.5 Self-supervised machine learning ... 23
7.6 Reinforcement machine learning ... 23
7.7 Transfer learning ... 24

8 Machine learning pipeline ... **25**
8.1 General ... 25
8.2 Data acquisition ... 26
8.3 Data preparation ... 27
8.4 Modelling ... 28
8.5 Verification and validation ... 30
8.6 Model deployment ... 30
8.7 Operation ... 30
8.8 Example machine learning process based on ML pipeline ... 31

Annex A (informative) **Example data flow and data use statements for supervised learning process** ... **34**

Bibliography ... **36**

© ISO/IEC 2022 – All rights reserved

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see patents.iec.ch).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 42, *Artificial Intelligence*.

Any feedback or questions on this document should be directed to the user’s national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

© ISO/IEC 2022 – All rights reserved

Introduction

Artificial intelligence (AI) systems, in general, are engineered systems that generate outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives. AI covers a wide range of technologies that reflect different approaches to dealing with these complex problems.

ML is a branch of AI that employs computational techniques to enable systems to learn from data or experiences. In other words, ML systems are developed through the optimisation of algorithms to fit to training data, or improve their performance based through maximizing a reward. ML methods include deep learning, which is also addressed in this document.

Terms such as knowledge, learning and decisions are used throughout the document. However, it is not the intent to anthropomorphize machine learning (ML).

This document aims to provide a framework for the description of AI systems that use ML. By establishing a common terminology and a common set of concepts for such systems, this document provides a basis for the clear explanation of the systems and various considerations that apply to their engineering and to their use. This document is intended for a wide audience including experts and non-practitioners. However, some of the clauses (identified in the overview in Clause 5), include more in-depth technical descriptions.

This document also provides the basis for other standards directed at specific aspects of ML systems and their components.

© ISO/IEC 2022 – All rights reserved

Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)

1 Scope

This document establishes an Artificial Intelligence (AI) and Machine Learning (ML) framework for describing a generic AI system using ML technology. The framework describes the system components and their functions in the AI ecosystem. This document is applicable to all types and sizes of organizations, including public and private companies, government entities, and not-for-profit organizations, that are implementing or using AI systems.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 22989, *Information technology—Artificial intelligence — Artificial intelligence concepts and terminology*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 22989 and the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1 Model development and use

3.1.1
classification model
<machine learning> machine learning model whose expected output for a given input is one or more classes

3.1.2
regression model
<machine learning> machine learning model whose expected output for a given input is a continuous variable

3.1.3
generalization
<machine learning> ability of a trained model to make correct predictions on previously unseen input data

Note 1 to entry: A machine learning model that generalizes well is one that has acceptable prediction accuracies using previously unseen input data.

Note 2 to entry: Generalization is closely related to overfitting. An overfit machine learning model will not generalize well as the model fits the training data too precisely.

© ISO/IEC 2022 – All rights reserved

3.1.4
overfitting
<machine learning> creating a model which fits the training data too precisely and fails to generalize on new data

Note 1 to entry: Overfitting can occur because the trained model has learned from non-essential features in the training data (i.e. features that do not generalize to useful outputs), excessive noise in the training data (e.g. excessive number of outliers) or because the model is too complex for the training data.

Note 2 to entry: Overfitting can be identified when there is a significant difference between errors measured on training data and on separate test and validation data. The performance of overfitted models is especially impacted when there is a significant mismatch between training data and production data.

3.1.5
underfitting
<machine learning> creating a model that does not fit the training data closely enough and produces incorrect predictions on new data

Note 1 to entry: Underfitting can occur when features are poorly selected, insufficient training time or when the model is too simple to learn from large training data due to limited model capacity (i.e. expressive power).

3.2 Tools

3.2.1
backpropagation
neural network training method that uses the error at the output layer to adjust and optimise the weights for the connections from the successive previous layers

3.2.2
learning rate
step size for a gradient method

Note 1 to entry: Learning rate determines whether and how fast a model converges to an optimal solution, making it an important hyperparameter to set for neural networks.

3.3 Data

3.3.1
class
human-defined category of elements that are part of the dataset and that share common attributes

EXAMPLE "telephone", "table", "chair", "ball bearing" and "tennis ball" are classes. The "table" class includes: a work table, a dining table, a study desk, a coffee table, a workbench.

Note 1 to entry: Classes are typically target variables and designated by a name.

3.3.2
cluster
automatically induced category of elements that are part of the dataset and that share common attributes

Note 1 to entry: Clusters do not necessarily have a name.

3.3.3
feature
<machine learning> measurable property of an object or event with respect to a set of characteristics

Note 1 to entry: Features play a role in training and prediction.

Note 2 to entry: Features provide a machine-readable way to describe the relevant objects. As the algorithm will not go back to the objects or events themselves, feature representations are designed to contain all useful information.

© ISO/IEC 2022 – All rights reserved

3.3.4
distance
<machine learning> measured proximity of two points in space

Note 1 to entry: Euclidean, or straight-line, distance is ordinarily used in machine learning.

3.3.5
unlabelled
property of a sample that does not include a target variable

4 Abbreviated terms

AI	artificial intelligence
API	application programming interface
AUC	area under the curve
BM	Boltzmann machines
CapsNet	capsule neural network
CG	conjugate gradient
CNN	convolutional neural network
DBN	deep belief networks
DCNN	deep convolutional neural network
FFNN	feed forward neural network
FNR	false negative rate
FPR	false positive rate
GRU	gated recurrent unit
LSTM	long short-term memory
MAE	mean absolute error
MDP	Markov decision process
ML	machine learning
NN	neural network
NNEF	neural network exchange format
NPV	negative predictive value
ONNX	open neural network exchange
PCA	principal component analysis
PHI	personal or protected health information
PII	personally identifiable information
PPV	positive predictive value

© ISO/IEC 2022 – All rights reserved