TECHNICAL REPORT

ISO/IEC TR 29119-11

First edition
2020-11

Software and systems engineering — Software testing —

Part 11: Guidelines on the testing of AI-based systems

Reference number
ISO/IEC TR 29119-11:2020(E)

© ISO/IEC 2020



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2020

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

Contents

Page

Foreword ..... **v**

Introduction ..... **vi**

1 Scope **1**

2 Normative references **1**

3 Terms, definitions and abbreviated terms **1**
3.1 Terms and definitions 1
3.2 Abbreviated terms 10

4 Introduction to AI and testing **11**
4.1 Overview of AI and testing 11
4.2 Artificial intelligence (AI) 11
4.2.1 Definition of 'artificial intelligence' 11
4.2.2 AI use cases 12
4.2.3 AI usage and market 12
4.2.4 AI technologies 13
4.2.5 AI hardware 15
4.2.6 AI development frameworks 16
4.2.7 Narrow vs general AI 16
4.3 Testing of AI-based systems 16
4.3.1 The importance of testing for AI-based systems 16
4.3.2 Safety-related AI-based systems 17
4.3.3 Standardization and AI 17

5 AI system characteristics **19**
5.1 AI-specific characteristics 19
5.1.1 General 19
5.1.2 Flexibility and adaptability 20
5.1.3 Autonomy 20
5.1.4 Evolution 21
5.1.5 Bias 21
5.1.6 Complexity 21
5.1.7 Transparency, interpretability and explainability 22
5.1.8 Non-determinism 22
5.2 Aligning AI-based systems with human values 23
5.3 Side-effects 23
5.4 Reward hacking 24
5.5 Specifying ethical requirements for AI-based systems 24

6 Introduction to the testing of AI-based systems **25**
6.1 Challenges in testing AI-based systems 25
6.1.1 Introduction to challenges testing AI-based systems 25
6.1.2 System specifications 25
6.1.3 Test input data 25
6.1.4 Self-learning systems 26
6.1.5 Flexibility and adaptability 26
6.1.6 Autonomy 26
6.1.7 Evolution 26
6.1.8 Bias 26
6.1.9 Transparency, interpretability and explainability 27
6.1.10 Complexity 27
6.1.11 Probabilistic and non-deterministic systems 27
6.1.12 The test oracle problem for AI-based systems 27
6.2 Testing AI-based systems across the life cycle 27
6.2.1 General 27
6.2.2 Unit/component testing 28

© ISO/IEC 2020 – All rights reserved

6.2.3 Integration testing 28
6.2.4 System testing 28
6.2.5 System integration testing 29
6.2.6 Acceptance testing 29
6.2.7 Maintenance testing 29

7 Testing and QA of ML systems 29
7.1 Introduction to the testing and QA of ML systems 29
7.2 Review of ML workflow 29
7.3 Acceptance criteria 29
7.4 Framework, algorithm/model and hyperparameter selection 30
7.5 Training data quality 30
7.6 Test data quality 30
7.7 Model updates 30
7.8 Adversarial examples and testing 30
7.9 Benchmarks for machine learning 31

8 Black-box testing of AI-based systems 31
8.1 Combinatorial testing 31
8.2 Back-to-back testing 32
8.3 A/B testing 32
8.4 Metamorphic testing 33
8.5 Exploratory testing 34

9 White-box testing of neural networks 34
9.1 Structure of a neural network 34
9.2 Test coverage measures for neural networks 36
9.2.1 Introduction to test coverage levels 36
9.2.2 Neuron coverage 36
9.2.3 Threshold coverage 36
9.2.4 Sign change coverage 36
9.2.5 Value change coverage 36
9.2.6 Sign-sign coverage 36
9.2.7 Layer coverage 37
9.3 Test effectiveness of the white-box measures 37
9.4 White-box testing tools for neural networks 37

10 Test environments for AI-based systems 38
10.1 Test environments for AI-based systems 38
10.2 Test scenario derivation 39
10.3 Regulatory test scenarios and test environments 39

Annex A Machine learning 40

Bibliography 49

 © ISO/IEC 2020 – 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).

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.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 7, *Software and systems engineering*.

A list of all parts in the ISO/IEC/IEEE 29119 series can be found on the ISO website.

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.

© ISO/IEC 2020 – All rights reserved

Introduction

The testing of traditional systems is well-understood, but AI-based systems, which are becoming more prevalent and critical to our daily lives, introduce new challenges. This document has been created to introduce AI-based systems and provide guidelines on how they might be tested.

Annex A provides an introduction to machine learning.

This document is primarily provided for those testers who are new to AI-based systems, but it can also be useful for more experienced testers and other stakeholders working on the development and testing of AI-based systems.

As a Technical Report, this document contains data of a different kind from that normally published as an International Standard or Technical Specification, such as data on the “state of the art”.

© ISO/IEC 2020 – All rights reserved

Software and systems engineering — Software testing —

Part 11: Guidelines on the testing of AI-based systems

1 Scope

This document provides an introduction to AI-based systems. These systems are typically complex (e.g. deep neural nets), are sometimes based on big data, can be poorly specified and can be non-deterministic, which creates new challenges and opportunities for testing them.

This document explains those characteristics which are specific to AI-based systems and explains the corresponding difficulties of specifying the acceptance criteria for such systems.

This document presents the challenges of testing AI-based systems, the main challenge being the test oracle problem, whereby testers find it difficult to determine expected results for testing and therefore whether tests have passed or failed. It covers testing of these systems across the life cycle and gives guidelines on how AI-based systems in general can be tested using black-box approaches and introduces white-box testing specifically for neural networks. It describes options for the test environments and test scenarios used for testing AI-based systems.

In this document an AI-based system is a system that includes at least one AI component.

2 Normative references

There are no normative references in this document.

3 Terms, definitions and abbreviated terms

3.1 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

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

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

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

3.1.1
A/B testing
split-run testing
statistical testing approach that allows testers to determine which of two systems or components performs better

3.1.2
accuracy
<*machine learning* (3.1.43)> performance metric used to evaluate a *classifier* (3.1.21), which measures the proportion of *classifications* (3.1.20) *predictions* (3.1.56) that were correct

© ISO/IEC 2020 – All rights reserved

3.1.3
activation function
transfer function
<*neural network* (3.1.48)> formula associated with a node in a neural network that determines the output of the node (*activation value* (3.1.4)) from the inputs to the neuron

3.1.4
activation value
<*neural network* (3.1.48)> output of an *activation function* (3.1.3) of a node in a neural network

3.1.5
adaptability
ability of a system to react to changes in its environment in order to continue meeting both functional and non-functional requirements

3.1.6
adversarial attack
deliberate use of *adversarial examples* (3.1.7) to cause a *ML model* (3.1.46) to fail

Note 1 to entry: Typically targets ML models in the form of a *neural network* (3.1.48).

3.1.7
adversarial example
input to an *ML model* (3.1.46) created by applying small perturbations to a working example that results in the model outputting an incorrect result with high confidence

Note 1 to entry: Typically applies to ML models in the form of a *neural network* (3.1.48).

3.1.8
adversarial testing
testing approach based on the attempted creation and execution of *adversarial examples* (3.1.7) to identify defects in an *ML model* (3.1.46)

Note 1 to entry: Typically applied to ML models in the form of a *neural network* (3.1.48).

3.1.9
AI-based system
system including one or more components implementing *AI* (3.1.13)

3.1.10
AI effect
situation when a previously labelled *AI* (3.1.13) system is no longer considered to be AI as technology advances

3.1.11
AI quality metamodel
metamodel intended to ensure the quality of *AI-based systems* (3.1.9)

Note 1 to entry: This metamodel is defined in detail in DIN SPEC 92001.

3.1.12
algorithm
ML algorithm
<*machine learning* (3.1.43)> algorithm used to create an *ML model* (3.1.46) from the *training data* (3.1.80)

EXAMPLE ML algorithms include linear regression, logistic regression, *decision tree* (3.1.25), SVM, Naive Bayes, kNN, K-means and random forest.

© ISO/IEC 2020 – All rights reserved