DRAFT
International Standard

ISO/IEC
DIS
24029-3

Artificial intelligence (AI) — Assessment of the robustness of neural networks —

Part 3: Methodology for the use of statistical methods

ICS: 35.020

ISO/IEC JTC **1**/SC **42**

Secretariat: **ANSI**

Voting begins on:
2026-02-02

Voting terminates on:
2026-04-27

THIS DOCUMENT IS A DRAFT CIRCULATED FOR COMMENTS AND APPROVAL. IT IS THEREFORE SUBJECT TO CHANGE AND MAY NOT BE REFERRED TO AS AN INTERNATIONAL STANDARD UNTIL PUBLISHED AS SUCH.

IN ADDITION TO THEIR EVALUATION AS BEING ACCEPTABLE FOR INDUSTRIAL, TECHNOLOGICAL, COMMERCIAL AND USER PURPOSES, DRAFT INTERNATIONAL STANDARDS MAY ON OCCASION HAVE TO BE CONSIDERED IN THE LIGHT OF THEIR POTENTIAL TO BECOME STANDARDS TO WHICH REFERENCE MAY BE MADE IN NATIONAL REGULATIONS.

RECIPIENTS OF THIS DRAFT ARE INVITED TO SUBMIT, WITH THEIR COMMENTS, NOTIFICATION OF ANY RELEVANT PATENT RIGHTS OF WHICH THEY ARE AWARE AND TO PROVIDE SUPPORTING DOCUMENTATION.

This document is circulated as received from the committee secretariat.

Reference number
ISO/IEC DIS 24029-3:2026(en)

© ISO/IEC 2026



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2026

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 and definitions **1**
4 Abbreviated terms **1**
5 Robustness assessment during the life cycle **2**
5.1 Summary 2
5.2 Inception 3
5.3 Design and development 3
5.4 Verification and validation 3
5.5 Deployment 4
5.6 Operation and monitoring 4
5.7 Continuous validation 4
5.8 Re-evaluation 5
5.9 Retirement 5
6 Robustness properties measurable using statistical methods **5**
6.1 General 5
6.2 In-domain robustness assessment 7
6.3 Specification of expected behaviour 7
6.3.1 General 7
6.3.2 Stability 8
6.3.3 Sensitivity 8
6.3.4 Relevance 8
6.3.5 Reachability 8
6.4 Other considerations that contribute to robustness assessment 8
7 Statistical measures of robust behaviour **9**
7.1 General 9
7.2 Robustness assessment based on performance drops 9
7.2.1 General 9
7.2.2 Examples 9
7.2.3 Thresholding process 9
7.3 Targeted assessment of the system's response to a perturbation 10
7.3.1 General 10
7.3.2 Examples 10
7.3.3 Thresholding process 10
7.4 Assessment through separability metrics 10
7.4.1 General 10
7.4.2 Example 11
7.4.3 Thresholding process 11
7.5 Assessment through statistical reliability engineering 11
7.5.1 General 11
7.5.2 Example 11
7.5.3 Thresholding process 12
7.6 Robustness assessment through curvature analysis 12
7.6.1 General 12
7.6.2 Examples 12
7.6.3 Thresholding process 12
8 Selection of datasets representative of a robustness dimension **12**
8.1 General 12
8.2 Datasets needed for robustness assessment 12
8.3 Acquisition of data for robustness assessment 13

© ISO/IEC 2026 – All rights reserved

8.4 Measuring adversarial robustness 14
Bibliography **15**

© ISO/IEC 2026 – All rights reserved

Foreword

ISO (the International *Organization* for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

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 ISO documents 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 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).

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

For an explanation on 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 the following URL: www.iso.org/iso/foreword.html.

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

A list of all parts in the ISO/IEC 24029 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 2026 – All rights reserved

Introduction

Neural networks are a widely used AI systems capable of tackling a variety of tasks, such as detection, classification or interpolation, on diverse data types, such as text, speech, image, time series or tabular data. AI quality models of neural networks comprise certain characteristics, including robustness. For example, ISO/IEC 25059:2023 [1] considers in its quality model that robustness is a sub-characteristic of reliability. Assessing robustness is one key step to build trust into AI systems, and in particular into neural networks. This document focuses on AI systems that comprise neural networks.

Demonstrating the ability of a system to maintain its level of performance under locally varying conditions can be done using formal methods, as illustrated in ISO/IEC 24029-2:2023 [2]. Statistical methods can be complementary to other methods in their ability to explore more widely these varying conditions.

As neural networks can be used to tackle high dimensionality data (e.g. real-world perception data, Internet of Things data), the input space they have to handle can be both variable and hard to explore exhaustively. In that regard statistical methods allow a wide exploration of their input space, which can be combined with any robustness assessment scheme to reinforce it.

© ISO/IEC 2026 – All rights reserved

Artificial intelligence (AI) — Assessment of the robustness of neural networks —

Part 3:
Methodology for the use of statistical methods

1 Scope

This document provides methodology for the use of statistical methods to assess robustness properties of neural networks.

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 22989:2022, *Information technology — Artificial intelligence — Artificial intelligence concepts and terminology*

ISO/IEC 23053:2022, *Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)*

ISO/IEC 24029-2:2023, *Artificial intelligence (AI) — Assessment of the robustness of neural networks — Part 2: Methodology for the use of formal methods*

3 Terms and definitions

For the purposes of this document, the *following terms and definitions given in ISO/IEC 22989:2022, as well as the following ISO/IEC 23053:2022* apply.

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

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

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

3.1
Stability
ability of an AI system to produce invariant output under specified perturbations on input

4 Abbreviated terms

AI	artificial intelligence
OOD	out-of-distribution

© ISO/IEC 2026 – All rights reserved