

ICS 35.100.05
CCS L79



**NATIONAL STANDARD OF THE
PEOPLE'S REPUBLIC OF CHINA**

GB/T 45225-2025

Artificial intelligence - Deep learning algorithms evaluation

人工智能 深度学习算法评估

Issued on: January 24, 2025

Implemented on: January 24, 2025

**Issued by: State Administration for Market Regulation;
Standardization Administration of the PRC**

Table of Contents

1 Scope	1
2 Normative references	1
3 Terms and Definitions	1
4 Evaluation Index System	2
5 Assessment Level	7
24 Reference	26

Foreword

This document is in accordance with the provisions of GB/T 1.1-2020 "Guidelines for standardization work Part

1. Structure and drafting rules for standardization documents" Drafting. Please note that some of the contents of this document may involve patents. The issuing organization of this document does not assume the responsibility for identifying patents. This document was proposed and coordinated by the National Technical Committee for Information Technology Standardization (SAC/TC28). The drafting units of this document are: China Electronics Technology Standardization Institute, Institute of Software, Chinese Academy of Sciences, Nanjing Software Technology Research Institute, Institute of Aeronautics and Astronautics, Beijing University of Aeronautics and Astronautics, Beijing Software Product Quality Inspection and Testing Center Co., Ltd., Beijing Institute of Aerospace Automatic Control, China South Power Grid Co., Ltd. Ultra-high Voltage Transmission Company, Shanghai Computer Software Technology Development Center, University of Science and Technology of China, Beijing Eyesight Technology Co., Ltd., Shanghai SenseTime Intelligent Technology Co., Ltd., Denso Intelligent Technology (Shanghai) Co., Ltd., China Electronics Technology Group Corporation Big Data Research Institute Co., Ltd., Inspur Electronic Information Industry Co., Ltd., China Mobile Communications Group Co., Ltd., Beijing Shengzhi Technology Co., Ltd., GRG Banking Group Co., Ltd., Shanghai Wenyu Information Technology Co., Ltd., Hangzhou Hikvision Digital Technology Co., Ltd., CASCO Signal Co., Ltd. Alibaba Cloud Computing Co., Ltd., Tianjin (Binhai) Artificial Intelligence Innovation Center, China Ordnance Industry Information Center, Shanghai Enflame Technology Co., Ltd. Co., Ltd., Shanghai Artificial Intelligence Industry Association, Shenzhen Yuntian Lifei Technology Co., Ltd., Sichuan Changhong Electronics Holding Group Co., Ltd., China Shipbuilding Industry Corporation Comprehensive Technical and Economic Research Institute, Beijing Institute of Computer Technology and Applications, Hong Kong University of Science and Technology, Chinese Academy of Sciences Institute

of Space Application Engineering and Technology, Zhejiang University, Shenyang Aircraft Design and Research Institute of Aviation Industry Corporation of China, Beijing University of Posts and Telecommunications, NARI Group Co., Ltd., Chongqing Guoke Chuzhi Information Technology Co., Ltd., Guoke Chushi (Chongqing) Software Co., Ltd., Chongqing Construction Industry (Group) Limited Liability Company. The main drafters of this document are. Bao Wei, Ye Heng, Meng Lingzhong, Xue Yunzhi, Ma Chenghao, Gao Hui, Liu Xianglong, Kong Hao, Wang Yang, Wang Ning, Chen Wenjie, Zhang Lan, Yang Chunlin, Wu Geng, Zhu Jian, Dong Qian, Yang Guang, Cai Huimin, Du Guoguang, Wang Kechen, Nie Jinran, Chen Xiaoliang, Xu Tianshi, Rui Ziwen, Ren Wenqi, Zhou Tingliang, Wu Tao, Shi Dianxi, Xie Wandong, Mei Jingqing, Chen Xi, Rao Xue, Cao Yu, Wu Lijin, Xu Zhewei, Song Jinke, Liu Aishan, Guo Jinyang, Wang Jinbo, Ji Shouling, Wen Xiaoling, Cheng Xiang, Chen Wei, Hu Yanling, Luo Yongjun, Zhang Yang. Artificial Intelligence Deep Learning Algorithm Evaluation

1 Scope

GB/T 45225-2025 sets the Chinese framework for evaluating deep learning algorithms. Its premise is that an algorithm cannot be assessed by accuracy alone: a model that scores well on a benchmark may be unusable because it is not robust to input perturbation, not reproducible, not explainable, too costly to run, or unsafe in ways the benchmark never tested. The standard therefore defines an evaluation indicator system spanning those dimensions, an assessment level scheme that grades an algorithm against them, and an evaluation process, the preparation, the selection of indicators and thresholds, the execution of the evaluation, the analysis of results and the report. Informative annexes give guidance on selecting the indicators and setting their thresholds for a given application, and worked material for the evaluation itself. It is one of the foundational documents of the Chinese AI standards programme, and for any vendor whose model will be assessed for use in China, or any organisation that must assess one, it is the framework the assessment will follow. It took effect on 24 January 2025.

This document establishes an evaluation index system for artificial intelligence deep learning algorithms and describes evaluation methods, etc. This document is intended to guide deep learning algorithm developers, users, third parties and other related organizations to deep learning algorithms and their training. The obtained deep learning model is used for evaluation.

2 Normative references

The contents of the following documents constitute the essential clauses of this document through normative references in this document. For referenced documents without a date, only the version corresponding to that date applies to this document; for referenced documents without a date, the latest version (including all amendments) applies to This

document.

GB/T 35273-2020 Information security technology Personal information security specification

GB/T 40660-2021 Information security technology Basic requirements for biometric information protection

GB/T 41867-2022 Information Technology Artificial Intelligence Terminology

3 Terms and definitions

The terms and definitions defined in GB/T 41867-2022 and the following apply to this document.

3.1 Deep learning A method for creating rich hierarchical representations by training neural networks with many hidden layers.

Note. Deep learning is a subset of machine learning. [Source: GB/T 41867-2022, 3.2.27]

3.2 A code snippet that uses a deep neural network structure to learn and reason to complete a specific function.

3.3 deep learning model deeplearningmodel A mathematical framework that produces inferences or predictions based on input data or information.

3.4 Test datatestdata Data used to evaluate the performance of the final machine learning model. [Source: GB/T 41867-2022, 3.2.3]

3.5 Adversarial examples Adding input samples formed by slight interference in the data set can induce the deep learning algorithm to give wrong output with a high probability, or even is to give a specific result.