ISO/IEC TS 42112:2026-06 (E)

Information technology - Artificial intelligence - Guidance on machine learning model training efficiency optimization

Contents

Page

Foreword iv

Introduction v

1 Scope 1

2 Normative references 1

3 Terms and definitions 1

4 Abbreviated terms 3

5 Overview of ML model training 3
5.1 Model training in ML pipeline 3
5.2 Stakeholders of model training 4

6 ML model training efficiency 4

7 Characteristics impacting ML model training efficiency 4
7.1 Training data 4
7.2 Model parameter management 4
7.3 Communication challenges 5
7.4 Failure detection in model training 5
7.5 Failure recovery in model training 5
7.6 Quality of the ML model 5
7.7 Management of computing resources 6

8 Model training efficiency optimization methods 6
8.1 Overview 6
8.2 Training data preparation and model optimization 6
8.2.1 General consideration 6
8.2.2 Training data quality optimization 7
8.2.3 Feature engineering 7
8.2.4 Feature selection 8
8.2.5 Feature scaling 8
8.2.6 Training algorithm selection 8
8.2.7 Training process optimization 9
8.2.8 Ensemble learning 9
8.3 Parallelism strategies 9
8.3.1 Data parallelism 9
8.3.2 Model parallelism 10
8.3.3 Hybrid parallelism 10
8.4 Communication optimization 10
8.4.1 Collective communication 10
8.4.2 Data compression 10
8.4.3 Asynchronous communication 11
8.4.4 Network topology-aware scheduling 11
8.5 Model checkpoint optimization 11
8.5.1 Hierarchical checkpoint saving 11
8.5.2 Overlapping model copy and computation 11
8.5.3 Network-aware asynchronous storage 11

8.6 Resource management for model training 12
8.7 Failure detection optimization 12
8.8 Continuous monitoring and anomaly detection 13
8.9 Environment and infrastructure assessment 13

Annex A (informative) Use case: Deep learning recommendation system for an e-commerce platform 14

Bibliography 17