ISO/IEC 5259-1:2024-07 (E)

Artificial intelligence - Data quality for analytics and machine learning (ML) - Part 1: Overview, terminology, and examples

Contents

Page

Foreword iv

Introduction v

1 Scope 1

2 Normative references 1

3 Terms and definitions 1

4 Symbols and abbreviated terms 5

5 Data quality concepts for analytics and machine learning 5
5.1 Data quality considerations for analytics and machine learning 5
5.1.1 General 5
5.1.2 Machine learning and data quality 5
5.1.3 Data characteristics that pose quality challenges for analytics and machine learning 6
5.1.4 Data sharing, data re-use and data quality for analytics and machine learning 6
5.2 Data quality concept framework for analytics and machine learning 6
5.2.1 Overview 6
5.2.2 Data quality management 7
5.2.3 Data quality governance 10
5.2.4 Data provenance 10
5.3 Data life cycle for analytics and ML 10
5.3.1 Overview 10
5.3.2 Data life cycle model 10
5.3.3 Processes across the multiple stages 13

Annex A (informative) Examples and scenarios 15

Bibliography 18