DIN SPEC 92001-2:2020-12 (E)

Artificial Intelligence - Life Cycle Processes and Quality Requirements - Part 2: Robustness

Contents

Page

Foreword ... 3

Introduction ... 5

1 Scope ... 6
1.1 Field of Application ... 6
1.2 Limitations ... 6

2 Normative References ... 7

3 Terms and Definitions ... 7
3.1 General Terminology ... 7
3.2 Terminology – Adversarial Robustness ... 7
3.3 Terminology – Corruption Robustness ... 9
3.3.1 Corruption Robustness – General terminology ... 9
3.3.2 Distributional Shift/ Dataset Shift ... 10
3.3.3 Sample Selection Bias ... 10

4 AI Quality Metamodel ... 11

5 Robustness ... 13
5.1 Introduction to AI Robustness ... 13
5.2 Requirements and Guidelines on Risk Management ... 14
5.2.1 Overview ... 14
5.2.2 Scope, Context, and Criteria ... 17
5.2.3 General Goals and Objectives ... 19
5.3 Requirements specific to Adversarial Robustness ... 20
5.3.1 Scope, Context, and Criteria ... 20
5.3.2 Threat Model Analysis ... 20
5.3.3 Likelihood & Impact Analysis ... 22
5.3.4 Robustness Evaluation ... 27
5.3.5 Mitigations ... 29
5.4 Requirements specific to Corruption Robustness ... 32
5.4.1 Scope, Context, and Criteria ... 32
5.4.2 Threat Model Analysis ... 32
5.4.3 Likelihood & Impact Analysis ... 34
5.4.4 Robustness Evaluation ... 37
5.4.5 Mitigations ... 39

6 Implementation Guidelines ... 41

Bibliography ... 43