Safety Integrity Levels for Artificial Intelligence

Simon Diemert[1], Laure Millet[1], Jonathan Groves[1], and Jeff Joyce[1]

1 Critical Systems Labs Inc., Vancouver, Canada

Abstract. Artificial Intelligence (AI) and Machine Learning (ML) technologies are rapidly being adopted to perform safety-related tasks in critical systems. These AI-based systems pose significant challenges, particularly regarding their assurance. Existing safety approaches defined in internationally recognized standards such as ISO 26262, DO-178C, UL 4600, EN 50126, and IEC 61508 do not provide detailed guidance on how to assure AI-based systems. For conventional (non-AI) systems, these standards adopt a 'Level of Rigor' (LoR) approach, where increasingly demanding engineering activities are required as risk associated with the system increases. This paper proposes an extension to existing LoR approaches, which considers the complexity of the task(s) being performed by an AI-based component. Complexity is assessed in terms of input entropy and output non-determinism, and then combined with the allocated Safety Integrity Level (SIL) to produce an AI-SIL. That AI-SIL may be used to identify appropriate measures and techniques for the development and verification of the system. The proposed extension is illustrated by examples from the automotive, aviation, and medical industries.

Accepted for presentation at the International Workshop on Artificial Intelligence Safety Engineering (WAISE) at SafeComp 2023, https://www.waise.org/programme