Bridging Gaps in ISO 26262 for Testing Machine Learning (ML) Systems

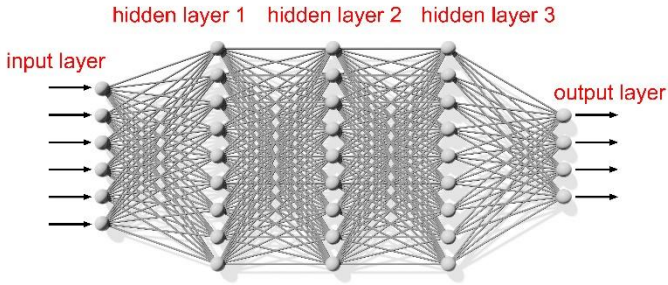


Integrating Machine Learning (ML) into safety-critical automotive systems presents new challenges. As automotive technologies evolve, standards like ISO 26262, which regulate the functional safety of electrical and electronic systems in vehicles, must be updated to address the complexities of modern ML systems. Drawing on a recent peer-reviewed IEEE Access publication co-authored with Emil Gracic and Gregor Pawelke, Padma Iyenghar outlines a systematic approach to enhancing ISO 26262 by introducing ML-specific life cycle phases and testing methods to ensure the safety and reliability of ML-driven systems.

While ISO 26262 [5] has effectively guided traditional software, it falls short in ensuring the safety and reliability of ML-driven systems due to the unique challenges posed by ML, such as non-determinism, interpretability issues, and the difficulty of comprehensive testing. As ML is increasingly embedded in automotive applications, adapting functional safety standards is crucial to address these challenges.

Gaps in ISO 26262 for ML

ISO 26262 is a solid standard and framework for ensuring the safety of automotive systems that includes structured guidelines for development, testing, and certification. However, the growing use of ML in safety-critical systems reveals significant gaps. These gaps mainly involve the life cycle phases of ML systems, their safety or desired properties, and the testing methods used to validate them.



deep neural network

1. Absence of ML-Specific Life Cycle Phases

While ISO 26262 covers software life cycle processes, it doesn't address the unique aspects of ML systems. ML involves stages like data preparation, model training, and deployment, each needing specific safety and verification measures. ISO 26262 does not guide these stages, leading to inconsistency and uncertainty in industry practices.

2. Lack of Safety Properties for ML Models

ML systems must meet certain safety properties that are not covered under traditional software safety requirements. For example, robustness, uncertainty handling, and interpretability are critical for ML models, particularly in scenarios where the systems interact with humans or make safety-critical decisions. These properties ensure that ML models operate safely under unpredictable conditions. However, ISO 26262 lacks provisions for their integration into the safety assurance process.

3. Insufficient Testing Methods

ISO 26262 provides specific recommendations for traditional software testing methods that vary based on the Automotive Safety Integrity Level (ASIL) assigned to each component or system. The rigour and depth of testing increase with each ASIL level, reflecting the higher risk and safety requirements.

"While existing standards like ISO/IEC 29119-11 and ISO PAS 8800 address some aspects of AI and ML testing, they do not comprehensively cover the full life cycle of ML models in automotive contexts"

This comprehensive approach ensures that systems meet their safety goals and function reliably under various conditions. However, ML testing introduces new challenges. ML models are often treated as "black boxes," complicating their interpretability, testing, and validation.

While existing standards like ISO/IEC 29119-11 [2] and ISO PAS 8800 [3] address some aspects of AI and ML testing, they do not comprehensively cover the full life cycle of ML models in automotive contexts. As a result, there is a lack of systematic, clear, and uniform development procedures that provide critical guidance for expert assessment and independent certification, ensuring reliable recommendations.

Proposed Framework: A systematic approach



To effectively integrate ML into the ISO 26262 framework for safety-critical automotive systems, our proposed solution in [1] outlines necessary enhancements that are required to bridge current gaps.

These enhancements are centred around the introduction of three new ML-specific life cycle phases: preparation of data, training of ML models, and deployment of ML models. Each of these phases necessitates defined inputs, rigorous testing methods, and the management of outputs that meet specified safety properties, such as robustness, explainability, and uncertainty handling. These properties are meticulously selected based on expert judgment to ensure they address the most critical aspects of ML system development.

ML-Specific Life Cycle Phases

To address the lack of ML-specific guidance in ISO 26262, three dedicated life cycle phases are proposed.

- **Preparation of Data:** This initial phase demands quality-assured data and metadata as inputs, which are critical for training robust ML models. The existing ISO 26262 framework does not specify these requirements; hence our proposal includes rigorous data handling measures to maintain the quality of ML-specific data.
- **Training of ML Model**: Post-data preparation, the ML model undergoes training. This phase's output often serves as a prototype and not the final deployed product. Thus, it's critical to apply stringent test methods during this phase to ensure the quality and reliability of the model before it advances to deployment.
- **Deployment of ML Model**: In the final phase, the trained ML model is refined and adapted to meet operational demands and hardware limitations, culminating in the deployment of the ML model. This version, which will be utilised in real-world applications, integrates with other software units and requires additional deployment-specific test methods to ensure its functionality and safety.

Definition of Desired Safety Properties

In the integration of ML into safety-critical automotive systems, defining desired properties for each life cycle phase is essential. These properties are critical non-functional requirements that ensure high-quality results throughout the development and testing of ML models. For instance, the trained ML model, a key outcome of one phase, becomes a foundational input for the deployment phase.

Selected based on expert judgment, desired properties such as completeness, consistency, robustness, explainability, and uncertainty handling are prioritised to ensure that ML models are fully developed and operate reliably under various conditions.

"The proposed framework aims to improve current practices in ML model development and testing, addressing gaps in existing standards"

By focusing on these key attributes, the proposed framework aims to improve current practices in ML model development and testing, addressing gaps in existing standards and laying the groundwork for the systematic testing of ML models, thereby contributing to the development of more reliable and safe automotive systems.

Definition of Test Methods

Several standards that provide robust test methods for developing and testing ML models were considered, each contributing valuable insights and methodologies. ISO 29119 (Part 11) and ISO PAS 8800 were selected based on expert judgment, recognising that many of these test methods are complementary and enhance their effectiveness in ensuring safety.

In the initial step, the test methods in these standards were thoroughly analysed for their applicability to the ML life cycle phases. This analysis resulted in a detailed mapping of test methods to specific phases of the ML life cycle. For each test method, a concise definition was developed, and a concrete and pragmatic description was provided to ensure clarity and ease of implementation.

It is important to emphasise that while specification and design are crucial activities within the life cycle phases discussed, the primary focus was placed on refining testing methodologies. These methods are essential in ensuring the robustness, reliability, and safety of ML models. By leveraging the comprehensive guidelines and methodologies from ISO 29119-11 and ISO PAS 8800, a solid foundation for the systematic testing of ML models was aimed to be established, ultimately contributing to the advancement of safety-critical automotive systems.

The systematic framework presented in [1] employs a rigorous evaluation framework inspired by IEC 61508 (Part 3, Annex C) [4], which assesses the effectiveness of methods for ASIL A and B conformance. This includes assigning rigour levels (R1 and R2) to categorise methods, ensuring that ML models meet essential safety properties, thereby enhancing product safety. The framework involves creating mapping tables that align desired safety properties with specific test methods, promoting a structured approach to their application.

Outputs and Recommendations

The framework establishes well-defined inputs, methods, and objectives for each life cycle phase, mirroring the structure used in ISO 26262. Through rigorous evaluations, it provides recommendations for methods that achieve ASIL A/B compliance. These enhancements aim to refine the testing and certification processes for ML-driven systems in the automotive industry, significantly boosting their safety and reliability.

Impact on the Industry

The proposed extension of ISO 26262 will have a significant impact on the automotive industry, particularly in the integration of ML into safety-critical systems. By defining clear life cycle phases and testing methods, the framework will:

- Provide clarity on ML model development and testing in compliance with functional safety requirements
- Streamline the certification process with a well-defined testing methodology
- Enhance the safety and reliability of ML systems through rigorous testing and alignment with key safety properties

This approach will be invaluable to automotive manufacturers, suppliers, and certification bodies, as it addresses the challenges of integrating ML into safety-critical systems while reducing ambiguity and enhancing overall system integrity.

Conclusion

The integration of ML into automotive systems is becoming increasingly prevalent, making it imperative to evolve ISO 26262 to address these modern complexities. Our proposed enhancements help bridge the current gaps, contributing to safer, more reliable ML-driven automotive systems. Future work will aim to extend this framework to higher ASIL levels and explore additional ML techniques, ensuring that safety standards keep pace with technological advancements.

For more detailed information and insights into our methodology and findings, please refer to our publication in [1].

References

[1] P. Iyenghar, E. Gracic and G. Pawelke, "A Systematic Approach to Enhancing ISO 26262 With Machine Learning-Specific Life Cycle Phases and Testing Methods," in IEEE Access, vol. 12, pp. 179600-179627, 2024, doi: 10.1109/ACCESS.2024.3506333.

[2] ISO/IEC TR, 29119-11:2020, "Software and Systems Engineering—Software Testing—Part 11: Guidelines on the Testing of AI-based Systems", 2020.

[3] ISO Std. ISO/DPAS, 8800, "Road Vehicles—Safety and AI", 2023.

[4] IEC 61508-1:2010, "Functional Safety of Electrical/electronic/programmable Electronic Safety Related Systems", 2010.

[5] ISO, 26262, "Road Vehicles—Functional Safety", 2018.

Image attribution:
Front image: AI generate by Midjourney
neural network: ID 105159234 © Eugenesergeev | Dreamstime.com
infographic: adapted from ID 210422761 © Macrovector Art | Dreamstime.com

Dr. Padma Iyenghar (innotec GmbH)



Padma Iyenghar is a senior Functional Safety and Cybersecurity Consultant at innotec GmbH–TÜV Austria Group, with a strong background in AI, academia, and industry. From 2020 to 2024, she also served as Development Professor at the University of Applied Sciences, Osnabrück, and continues to take on selected teaching assignments. She holds a Ph.D. from the University of Osnabrück and has over 14 years of R&D and 8 years of teaching experience, contributing actively to academic and professional communities. She is also an active member of the standards committee DKE/AK 914.0.11 – Functional Safety and AI.

Web: https://sites.google.com/site/piyengha/