On the Safety of Automotive Systems Incorporating Machine Learning based Components A position paper

Mohamad Gharib[1], Paolo Lollini[1], Marco Botta[2], Elvio Amparore[2], Susanna Donatelli[2] and Andrea Bondavalli[1]

[1]University of Florence, Florence, Italy
[2]University of Torino, Torino, Italy

The 4th Workshop on Safety and Security of Intelligent Vehicles (SSIV'18)
June 25, 2018

Outline

- The problem statement
- Research baseline
  - ISO 26262
  - Machine learning and safety
- Illustrative example: Maneuver Assistance System
- Formulating the Research Questions
- Proposed Solution
- Conclusions

The Problem Statement

- The last few years have witnessed an increasing adoption of Machine learning (ML) components in many automated systems for performing complex tasks such as pattern recognition, image recognition, and even control.
- Some of these systems can be classified as ***safety-critical systems***, where their failure may cause death or injury to humans.
- Therefore, the safe use of such systems should be ***assessed and assured*** before they are actually used in their operational environment.

The Problem Statement

- The last few years have witnessed an increasing adoption of Machine learning (ML) components in many automated systems for performing complex tasks such as pattern recognition, image recognition, and even control.
- Some of these systems can be classified as ***safety-critical systems***, where their failure may cause death or injury to humans.
- Therefore, the safe use of such systems should be ***assessed and assured*** before they are actually used in their operational environment.

There is neither a safety standard for certifying automotive systems that incorporate ML-based components, nor a concrete agreed upon method for their V&V.

Research Baseline

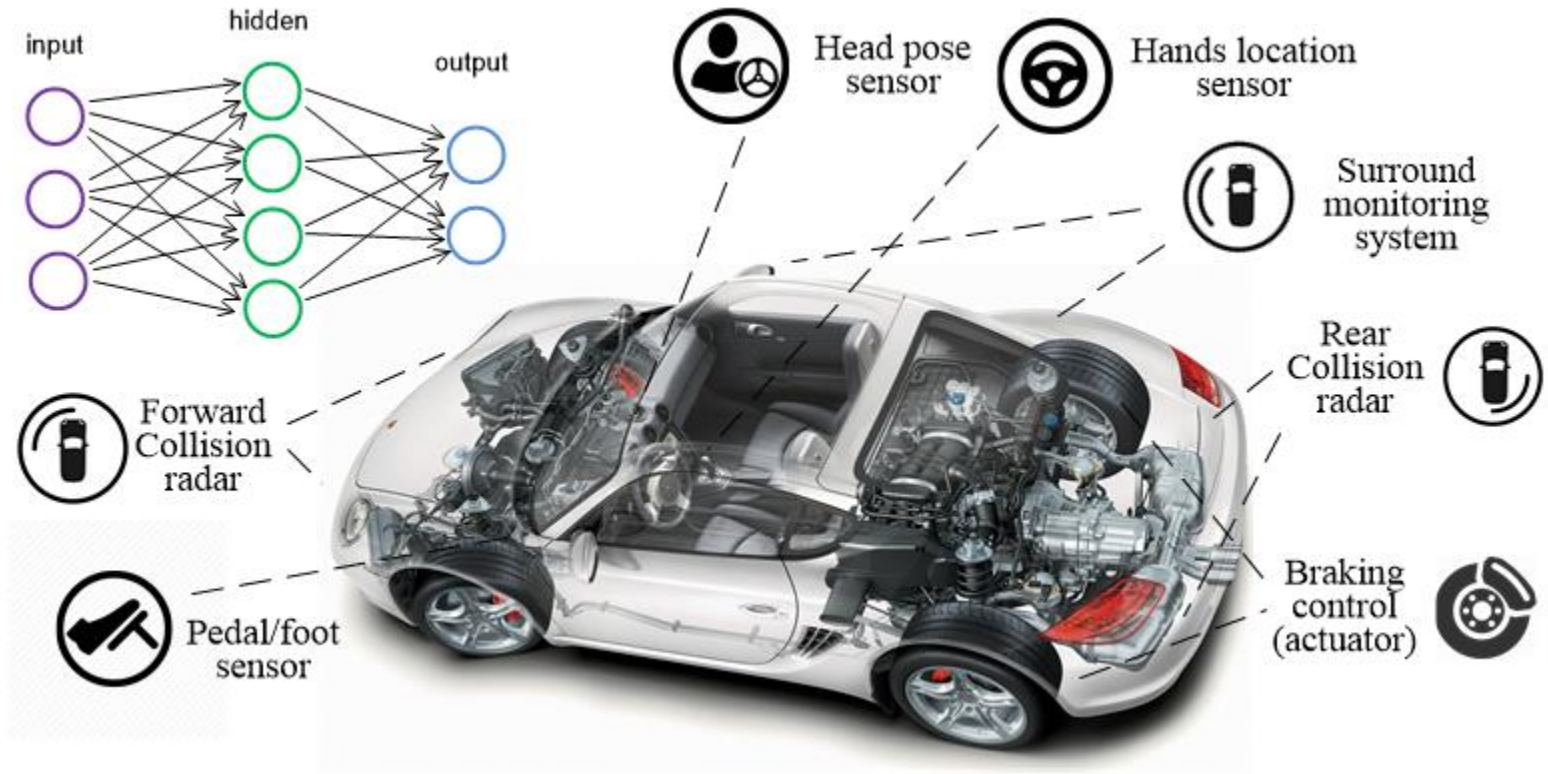
- ISO 26262
  - ISO 26262:2011 is a ***functional safety standard*** applicable to all road vehicles with a weight under 3500 kg.
  - ISO 26262 has been developed with a main objective to provide guidelines and best practices to increase the safety of Electronic and Electric (E/E) systems in vehicles.

Research Baseline

- ISO 26262
  - ISO 26262:2011 is a ***functional safety standard*** applicable to all road vehicles with a weight under 3500 kg.
  - ISO 26262 has been developed with a main objective to provide guidelines and best practices to increase the safety of Electronic and Electric (E/E) systems in vehicles.

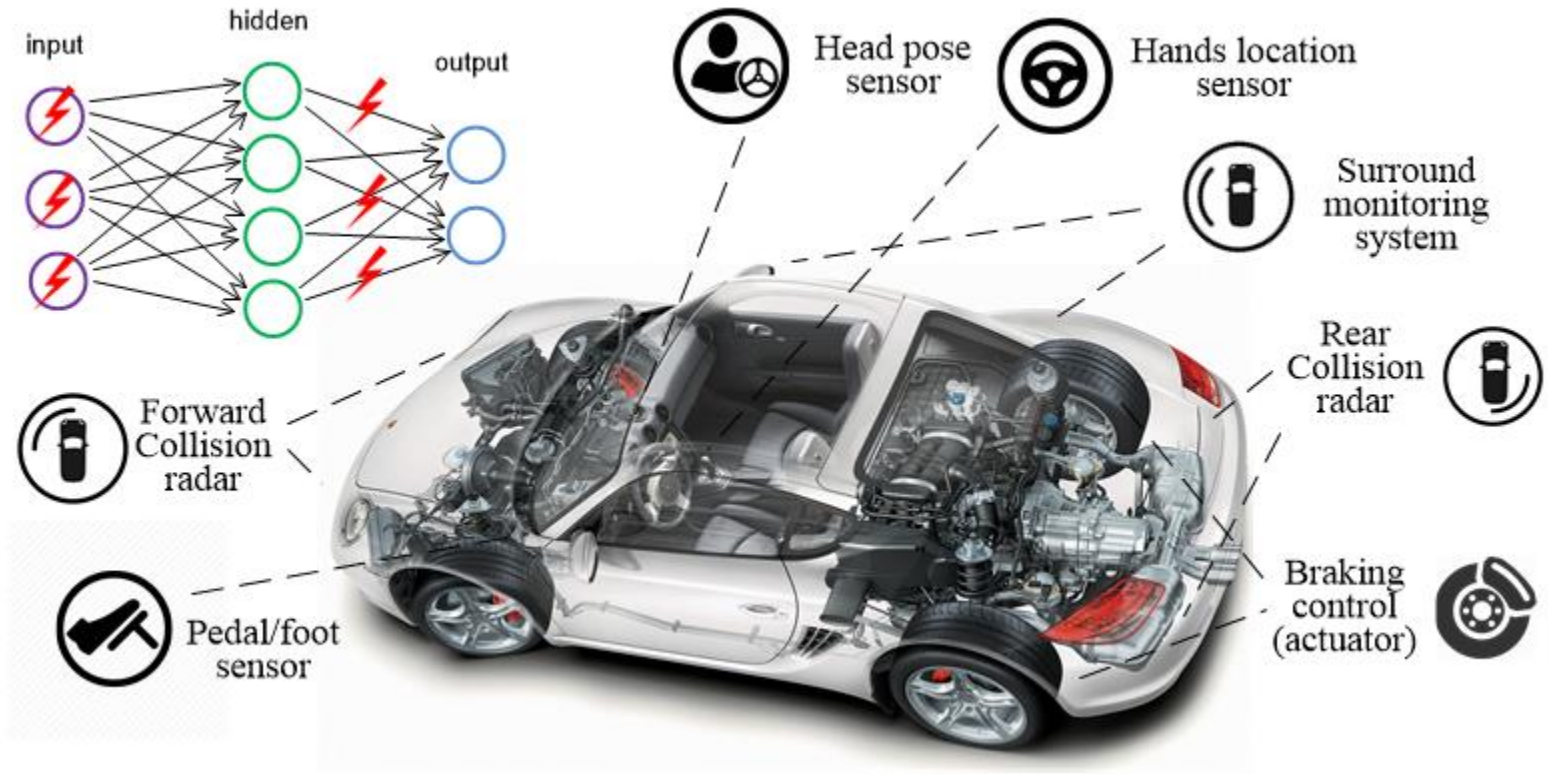
- Machine Learning and Safety
  - Artificial Neural Networks (ANN) that are organized into well-defined structures, such as discrete layers, grids, etc.
  - An ML algorithm makes ***predictions*** based on a model calculated from its input data. Inherently, these predictions carry some chance of being wrong, i.e., there is ***uncertainty*** that the ML algorithm will make a correct prediction.

Example: Maneuver Assistance System



- **The Maneuver Assistance System (MAS)** is expected to increase the driver's safety by monitoring its behaviour, detecting unintended maneuvers, and responds in a way that guarantees the highest possible level of driver safety.

Example: Maneuver Assistance System



- The ML-based component may ***wrongly*** categorize an intended/unintended maneuver as an unintended/intended one, which will prevent/allow performing an intended/unintended maneuver.

Both of these situations can be life-threatening

Formulating the Research Questions

- **RQ1:** How ML-based components are different from traditional software components with respect to safety-related aspects?

Formulating the Research Questions

- **RQ1:** How ML-based components are different from traditional software components with respect to safety-related aspects?
- **RQ2:** How the main safety-related concepts should be modified to find an adequate context to limit the safety risks of automotive systems that incorporate ML-based components?

Formulating the Research Questions

- **RQ1:** How ML-based components are different from traditional software components with respect to safety-related aspects?
- **RQ2:** How the main safety-related concepts should be modified to find an adequate context to limit the safety risks of automotive systems that incorporate ML-based components?
- **RQ3:** How can we assess the likelihood and severity of ML-based components errors, faults and failures?

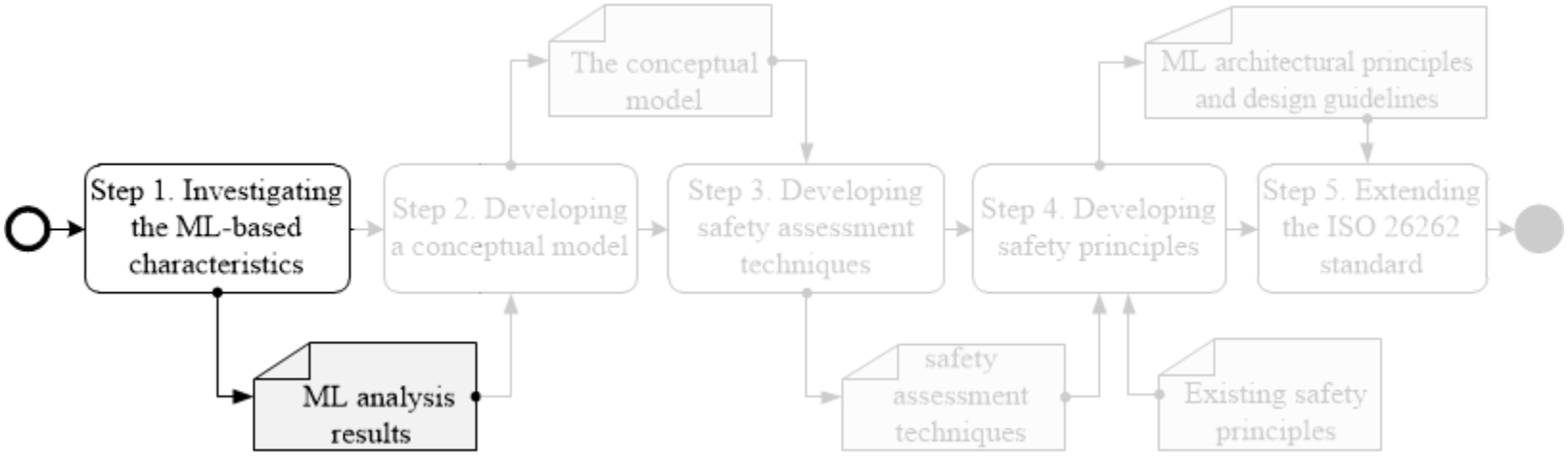
Formulating the Research Questions

- **RQ1:** How ML-based components are different from traditional software components with respect to safety-related aspects?
- **RQ2:** How the main safety-related concepts should be modified to find an adequate context to limit the safety risks of automotive systems that incorporate ML-based components?
- **RQ3:** How can we assess the likelihood and severity of ML-based components errors, faults and failures?
- **RQ4:** Do existing safety principles applies to the design of automotive systems that incorporate ML-based components?

Formulating the Research Questions

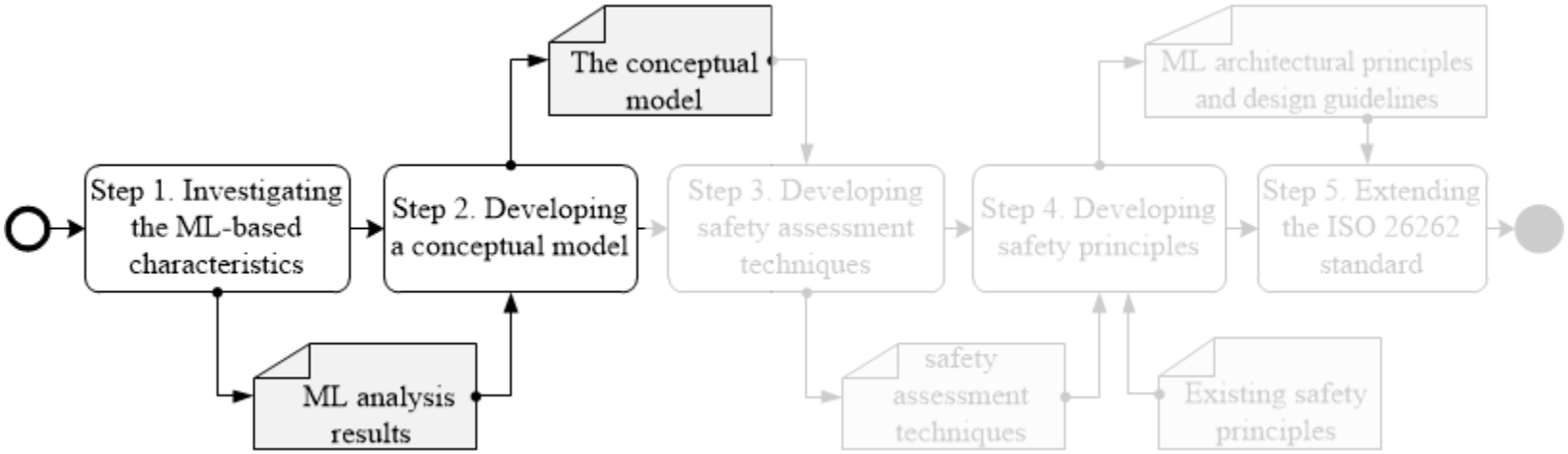
- **RQ1:** How ML-based components are different from traditional software components with respect to safety-related aspects?
- **RQ2:** How the main safety-related concepts should be modified to find an adequate context to limit the safety risks of automotive systems that incorporate ML-based components?
- **RQ3:** How can we assess the likelihood and severity of ML-based components errors, faults and failures?
- **RQ4:** Do existing safety principles applies to the design of automotive systems that incorporate ML-based components?
- **RQ5:** How the ISO 26262 standard should be extended to address the use of ML-based components?

Proposed Solution



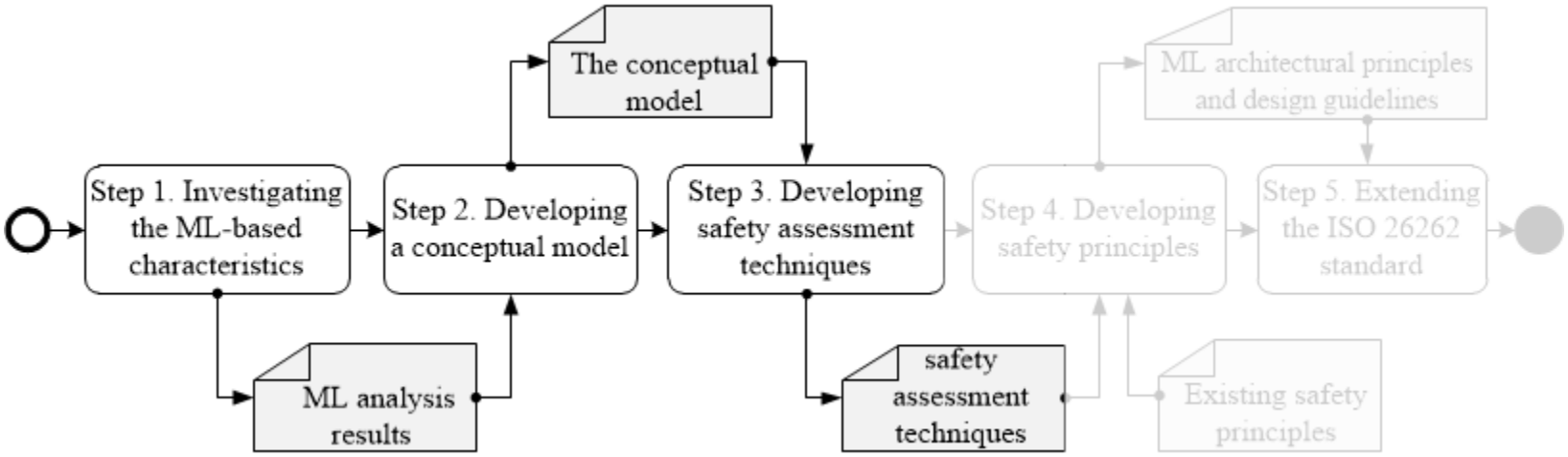
- **Step 1.** Investigating the specific characteristics of ML-based components with respect to safety-related aspects:
  - ML makes predictions based on a model calculated from its input data, and these predictions carry some chance of being wrong.
  - The specific characteristics of ML to be investigated are their non-determinism, non-transparency, their error rate, and their instability.

Proposed Solution



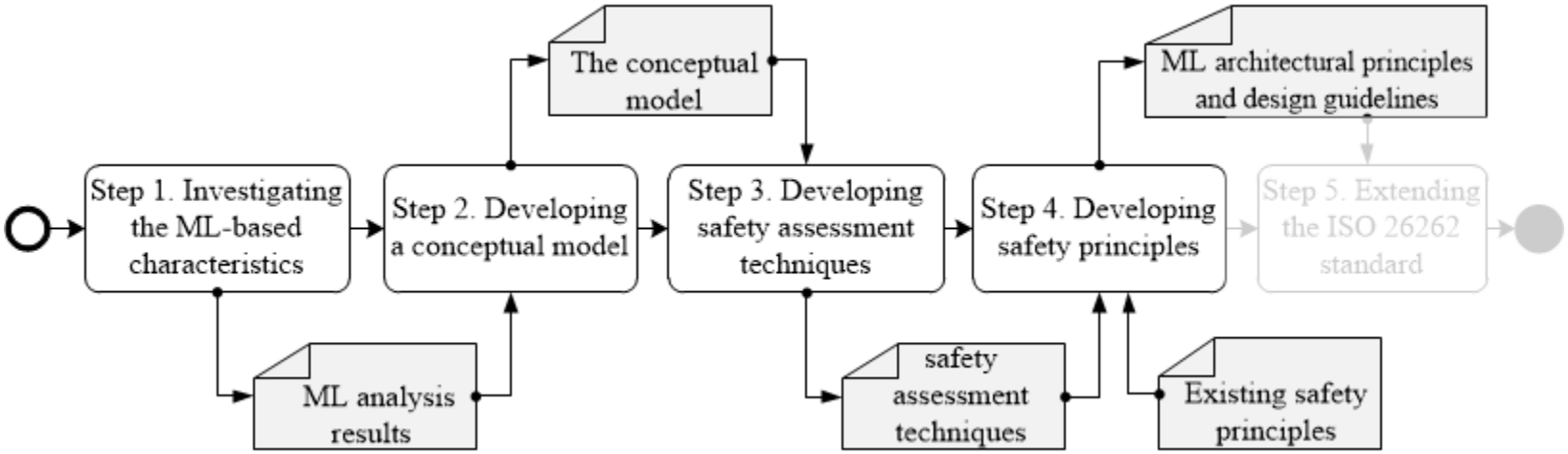
- **Step 2.** Developing a conceptual model for modeling automotive systems incorporating ML-based components:
  - We will develop a conceptual model that identifies the main safety concepts of both automotive systems and ML-based components, the different interrelations among these concepts, and how such concepts should be adapted/modified to be used for modeling automotive systems incorporating ML-based components.

Proposed Solution



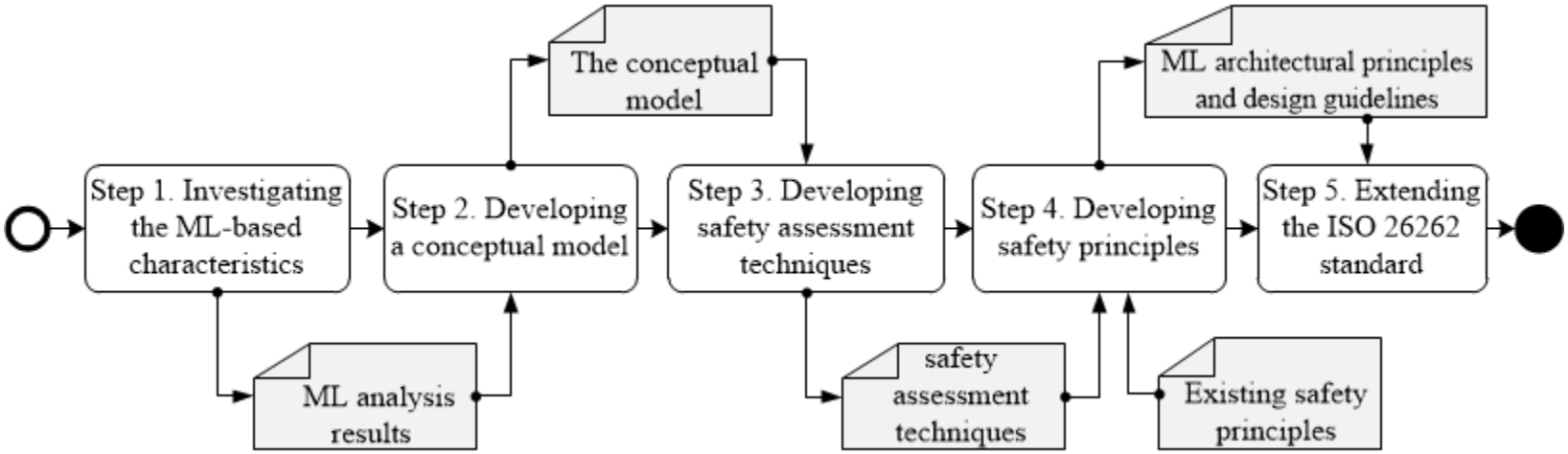
- **Step 3.** Developing safety assessment techniques:
  - We will develop techniques for assessing the safety of automotive systems incorporating ML-based components. These techniques will enable to detect the occurrence of ML errors, faults and failures as well as assessing their consequences, which can be used for avoiding, tolerating and/or mitigating any safety-related consequences that may arise due to such errors, faults and failures.

Proposed Solution



- **Step 4.** Developing safety principles for assuring the safe incorporation of ML-based components in automotive systems:
  - We will conduct an extensive analysis of existing safety engineering principles to develop architectural principles and design guidelines for assuring the safety of automotive systems incorporating ML-based components.

Proposed Solution



- **Step 5.** Extending ISO 26262 standard to address the use of ML-based components:
  - We will investigate how the safety standard ISO 26262 should be adapted/extended to consider the proposed solutions developed within this research.

Conclusion

- We argued that the safe use of ML-based components in automotive systems ***must be assessed and assured*** before they are used in their operational environment.
- Our argument is based on the ***ML and safety related literature***, in which we were not able to find neither a safety standard for certifying automotive systems that incorporate ML-based components, nor an agreed upon method for their V&V.
- We have presented our view on the safety of automotive systems incorporating ML-based components.
- We have motivated and provided ***a research agenda*** for extending the ISO 26262 standard to address challenges posed by incorporating ML-based components in automotive systems.

THANK YOU
for your attention