Artificial Intelligence (AI) for Safety-Critical Systems

Promising technologies for ECS
22nd September 2022

Jon Perez
jmperez@ikerlan.es
(INSIDE SC)
Example AI and dependable / Safety systems

- Operating conditions:
  - Closed World
  - Open World

- Level of intelligence:
  - Automation level:
    - Automatic
    - Heteronomous
    - Autonomous

- Dependability:
  - Reliability, availability
  - Safety & Safety Certification

- High Computing Performance (HPC)
Machine learning

Artificial intelligence

Machine learning

Deep learning

- Linear regression.
- Logistic regression.
- Bayesian networks.
- Decision tree.
- Greedy algorithm.
- Support Vector Machine (SVM).
- Naive bayes.
- Random forest.
- Gradient boosting algorithm.

- Neural Network (NN).
- Deep learning.
- Recurrent Neural Network (RNN).
- Long Short Term Memory (LSTM).

Artificial Intelligence [ISO 22989]:
“set of methods or automated entities that together build, optimize and apply a model so that the system can, for a given set of predefined tasks, compute predictions, recommendations, or decisions”

Machine Learning [Oxford]:
“the use and development of computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyze and draw inferences from patterns in data”
Scope

Updates

ML Development (Training)

ML Development (Training)

SAFETY STANDARDS

Trustworthiness
Safety standards

Functional Safety (Traditional)

Het./Autonomous Systems
Artificial intelligence

Drafts:
- **ISO/PAS 21448**: Road vehicles — Safety of the intended functionality (SOTIF).
- **UL 4600**: Safety for the evaluation of autonomous products.
- **ISO/IEC AWI TR 5469**: Artificial intelligence — functional safety and AI systems.
- **VDE-AR-E 2842-61-1**: Development and trustworthiness of autonomous/cognitive systems
- Etc.

Working Groups: EUROCA WG-114, SAE G-34, etc.
Scope

Updates

ML Development (Training)

Artificial intelligence, autonomous systems and safety (draft) standards

SOTIF, UL 4600, IEC 5469

Safety standards

Functional safety standards: electronics / programmable electronics

Trustworthiness

ML Development (Training)
Safety standards

AI Lifecycle
ML Lifecycle

Example safety case
[Bu17] – GSN

[AS19]
Development (training)
ML Development (Training)

Safety standards

SOTIF, UL 4600, IEC 5469
Artificial intelligence, autonomous systems and safety (draft) standards

Functional safety standards: electronics / programmable electronics

Trustworthiness

Updates

Scope
Deployment (inference)

**Hardware device**
- Multi-core devices and MPSoCs [Pe20]
- GPUs [Ko19, Le18, Pe22]
- FPGAs [Ko19]
- Specialized devices, e.g., TPU
- Proprietary devices, e.g. Tesla FSD [Ta20]

**Software**

High-Performance Embedded Computing (HPEC) – Hardware and software.
- Safety compliance of ‘Parallel Programming runtime’ and ‘ML software libraries’.
- Temporal Independence, guarantees and diagnosis [Pe20].
- Spatial independence [Pe20].
- Diagnosistic coverage (DC) [Pe20].
- Thermal dissipation, energy consumption [Ko19].
## Deployment (inference)

### Hardware device

- Multi-core devices and MPSSoCs [Pe20]
- GPUs [Ko19, Le18, Pe22]
- FPGAs [Ko19]
- Specialized devices, e.g., TPU
- Proprietary devices, e.g. Tesla FSD [Ta20]

---

High-Performance Embedded Computing (HPEC) – Hardware and software.

- Safety compliance of ‘Parallel Programming runtime’ and ‘ML software libraries’.
- Temporal Independence, guarantees and diagnosis [Pe20].
- Spatial independence [Pe20].
- Diagnosistic coverage (DC) [Pe20].
- Thermal dissipation, energy consumption [Ko19].

---

### Safety Std. Compliance and support

Co17

- Safety compliance of Parallel Programming runtime
- ML software libraries

Temporal Independence, guarantees and diagnosis [Pe20]

Spatial independence [Pe20]

Diagnosistic coverage (DC) [Pe20]

Thermal dissipation, energy consumption [Ko19]

### Specialized computing solutions

Co17

- NVIDIA Tesla FSD [Ta20]

---

### Generic purpose computing

- Multi-core devices and MPSSoCs
- GPUs [Ko19, Le18, Pe22]
- FPGAs [Ko19]
- Specialized devices, e.g., TPU
- Proprietary devices, e.g. Tesla FSD [Ta20]
Scope

Updates

ML Development (Training)

Safety standards

Trustworthiness

ML Development (Training)

SOTIF, UL 4600, IEC 5469
Artificial intelligence, autonomous systems and safety (draft) standards

Safety standards

Functional safety standards: electronics / programmable electronics

© 2022 IKERLAN. All rights reserved
Trustworthiness

Engineering

Ethics

Legal

Engineering Ethics

Machine Ethics

EUROPEAN COMMISSION

Brussels, 21 April 2021
COM(2021) 206 final
2021/0106(COD)

Proposal for a

REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL
LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS
There is a need to pave the way towards the development and certification of AI-based safety systems:

• Need for AI safety standard(s) definition and consolidation, complementary with functional safety standards.
• Many ML technical challenges: training data coverage (e.g., corner cases), understability, testability, verifiability, etc.
• Evolution of HPC challenges (HW / SW): integration of functions of different criticality in a HPC safe execution platform:
  • Safety compliance of ML libraries and parallel programming languages.
  • Temporal and spatial independence.
  • Thermal and energy requirements.
• Trustworthiness: Engineering, Ethics and Legal
• Safe and secure update of systems
THANK YOU

www.ikerlan.es

P.º J. M.ª. Arizmendiarieta, 2 - 20500 Arrasate-Mondragón
T. +34 943 712 400 F. +34 943 796 944