consorzio interuniversitario nazionale per l'informatica

UNIVERSITÀ DEGLI STUDI DI MILANO

IEEE WCNC

ML Assurance in 6G-Enabled Edge-Cloud Continuum Workflows

Alex Della Bruna, Filippo Berto, Marco Anisetti, Claudio A. Ardagna

24th March 2025

SESAR LAB

Machine Learning Models

Machine learning models are used in a wide range of applications

- **Image recognition**
- **Natural language processing**
- **Action automations**

These models are trained on large datasets and are used to make predictions on new data

CIA Triad

The CIA properties are the three core principles of information security that ensure that

- **Confidentiality** - access policies are enforced
- **Integrity** - information has not been tampered
- **Availability** - data are accessible when needed

This triad can be extended to include other properties such as authenticity, non-repudiation, and accountability

Assurance

"**Security assurance** is the process of **asserting non-functional properties** of a system to **ensure** that it meets the given **non-functional requirements** over time"[1]

[1] ***C. A. Ardagna and N. Bena.*** *"Non-Functional Certification of Modern Distributed Systems: A Research Manifesto".* ***In: Proc. of IEEE SSE 2023. Chicago, IL, USA, July 2023***

Assurance of Machine Learning Models

Goal: Assess machine learning models against CIA triad

- **Confidentiality** - Ensuring that the models and data are only shared with intended recipients
- **Integrity** - Ensuring that the models are consistent between releases
- **Availability** - Ensuring that the models respond within a certain amount of time

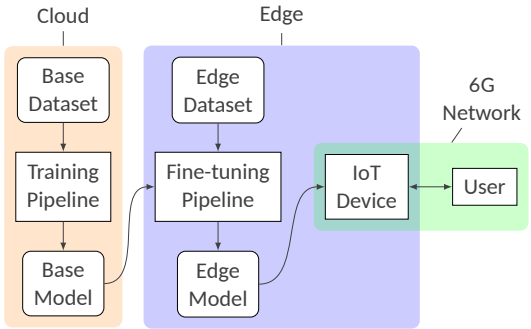
Assurance of Machine Learning Models: Challenges

- **Complexity** - Limited explainability and interpretability
- **Black-box nature** - Not transparent internal working
- **Context awareness** - Need to consider the external context

A Practical Scenario

Our working scenario entails three components

- **ML Model** - A machine learning image classification model[2]
- **6G Network** - A 6G-compliant network used to provide a local high performance deployment
- **Workflow** - A workflow that is used to train and deploy the model on the 6G network using fine-tuning techniques



[2]https://huggingface.co/microsoft/resnet-50

Assurance of Machine Learning Models: Integrity

Two approaches

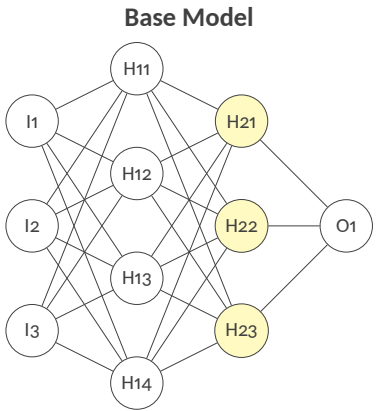
- **Structural Integrity** - evaluates the model's **structural differences** between different releases
- **Behavioral Integrity** - evaluates the model's **inference capacity** between different releases

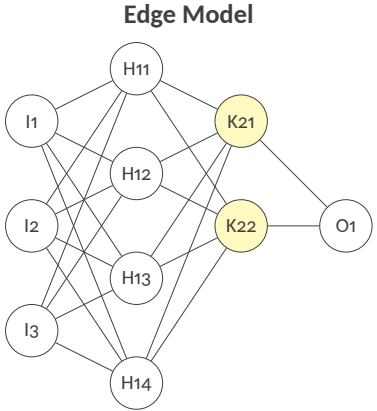
Structural Integrity

Evaluate the structural differences between releases of the model

- Use the **euclidean norm** to measure the differences between the **fully connected layers**
- Identify a **difference threshold**
- Define the property's **contract**

Structural Integrity 2





Structural Integrity 3

Starting from the definition of **euclidian norm**

$$\|A - B\|_2 = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij} - b_{ij})^2}$$

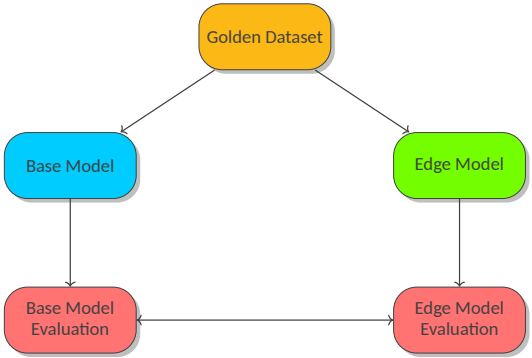
We have defined our **contract** as follows

$$\begin{aligned} c_{\text{struct. integ.}}(M_1, M_2) = \mathit{let} \\ L_1 = M_1.\texttt{FC\_layers[-1]} \\ L_2 = M_2.\texttt{FC\_layers[-1]} \\ \mathit{in}\ \|L_1 - L_2\| \leq \mathit{threshold} \end{aligned}$$

Behavioral Integrity

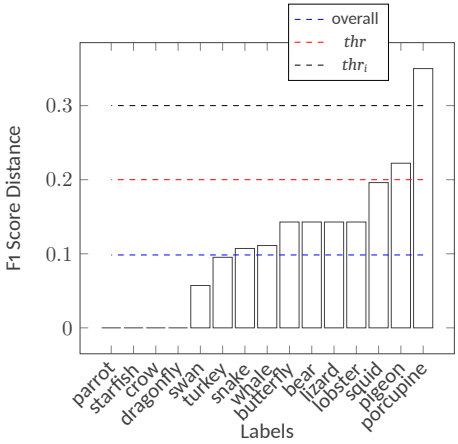
Evaluate the performance of the model in terms of **inference capacity** between different releases

- Use the **F1 score** to quantify the model's performance
- Identify **global and local performance thresholds**
- Define the property's **contract**



Behavioral Integrity 2

$$c_{\text{contx. integ.}}(M_1, M_2, S) = let$$
$$f1_{\text{tot}} = |F1(M_1, S) - F1(M_2, S)| \leq thr$$
$$f1_i = |F1(M_1, S_i) - F1(M_2, S_i)| \leq thr_i$$
$$in\, f1_{\text{tot}} \wedge \bigwedge_{i \in labels} f1_i$$



Conclusions

We have proposed a **novel approach for the assurance of machine learning models** that

- **Assesses the CIA triad** of machine learning models, focusing on the **integrity** property
- **Defines two different approaches** for evaluating the integrity of machine learning models, **structural integrity, and behavioral integrity**

Future works

- **Extend** the proposed approach **including other non-functional properties**
- **Evaluate** the proposed approach **on a wider range of machine learning models and workflows**

ML Assurance in 6G-Enabled Edge-Cloud Continuum Workflows

Thank you for listening!
Any questions?

Alex Della Bruna: alex.dellabruna@consorzio-cini.it
Filippo Berto: filippo.berto@unimi.it
Marco Anisetti: marco.anisetti@unimi.it
Claudio A. Ardagna: claudio.ardagna@unimi.it