Requirements regarding ONNX format to be used for ML model description coming from the ARP6983 draft5b

Reference document: ARP6983-ED324 Draft 5b

Author: JB Rouffet

Validator:

Authorizer:

The ML Model Description (MLMD) is an important artifact for transitioning the trained ML Model from the MLDL process to the implementation phase, ensuring its correct and safe integration into the system.

ARP6983 document outlines several key requirements for the ML Model Description (MLMD) to ensure it can be effectively used as an input for the implementation phase (second V of the Wshape process). These requirements are designed to ensure the MLMD provides a clear, complete, and unambiguous specification of the ML Model, facilitating its correct and safe implementation within the airborne system.

Here's a proposed breakdown of these requirements:

# 1. Architectural Clarity:

### ML Model Architecture

The MLMD must clearly describe the logical architecture of the ML Model. This includes identifying the individual ML Model elements (e.g., submodels, layers, nodes), their learning paradigm (e.g., supervised, unsupervised), the associated ML technology (e.g., neural networks, decision trees), and how these elements are interconnected to achieve the intended function(s). (See 6.4.3.6.a)

“Examples of ML Model architecture description content include (but are not limited to):

* The problem addressed by the proposed architecture (e.g., minimization, maximization, clustering, regression, classification, anomaly detection, optimal control),
* ML Model inputs (list, types, definition, units, accuracy, and resolution),
* ML Model output (list, types, definition, units, accuracy, and resolution),
* The logical structure of the proposed ML Model including the ML Model elements, their learning paradigm (e.g.supervised learning), the associated ML technology (e.g., neural networks), the function they perform in the architecture and their interfaces in terms of data flows and control flows
* Sequence constraints to be met (e.g., sequence diagram).”

### Interfaces:

The MLMD should define the interfaces between the ML Model architecture elements in terms of data flow and control flow, ensuring consistency and proper interaction between these elements. (See 6.4.3.6.a)

### Sequence Constraints:

If there are specific sequence constraints that need to be met during the execution of the ML Model (e.g., a specific order of operations for submodels), these constraints should be clearly documented in the MLMD, potentially using a sequence diagram. (See 6.4.3.6.a)

# 2. Parameter Specification:

### ML Model Hyperparameters:

The MLMD must specify the values of the ML Model hyperparameters. These hyperparameters are determined during the ML Model design process and control the learning process, ultimately influencing the values of the ML Model parameters. Examples include learning rate, number of hidden layers, and regularization parameters. (See 6.4.3.6.b)

### ML Model Parameters:

The MLMD should include a description of the ML Model parameters, which are internal to the ML Model and are estimated from data during the training process. Examples include weights in a neural network, support vectors in a support vector machine, and coefficients in a linear regression model. (See 6.4.3.6.c)

# 3. Algorithmic Transparency:

### Analytical/Algorithmic Syntax and Semantics:

The MLMD must provide a detailed and unambiguous description of the analytical/algorithmic form of the ML Model. This includes specifying the mathematical operations, algorithms, and internal logic used by the ML Model to compute its outputs from its inputs. This description should be sufficiently detailed to facilitate the implementation of the ML Model in software or hardware. (See 6.4.3.6.d)

# 4. Replication and Execution:

### Replication Criterion:

The MLMD should define the replication criterion for the ML Model, specifying whether exact or approximate replication is required. This criterion determines the level of detail required in the MLMD to ensure the implemented ML Model faithfully reproduces the behavior of the designed ML Model. (See 6.4.3.6.e)

### Execution Environment:

The MLMD must specify the execution environment of the ML Model, including the target hardware platform, operating system, libraries, and any other dependencies required for the ML Model to function correctly. (See 6.4.3.6.f)

# 5. Clarity and Completeness:

### Independence from Learning Environment:

The MLMD should be developed to avoid unnecessary adherence to the specific learning environment used during the ML Model design process. Any dependencies on the learning environment (e.g., specific libraries or formats) should be explicitly mentioned. (See 6.4.3.6.g)

### Removal of Extraneous Information:

The MLMD should not include any information that is not part of the implemented ML Model. This includes comments, illustrations, informative metrics, tool traces/logs, and other extraneous data that could lead to confusion or ambiguity during the implementation process. (See 6.4.3.6.h)

Some of those elements will have to be deepdived as the WG-SONNX has already identified some elements that could influence the execution of trained models (see numerical accuracy page: <https://github.com/ericjenn/working-groups/blob/ericjenn-srpwg-wg1/safety-related-profile/meetings/numerical_issues/2024-11-06%20-%20Numerical%20issues.pdf>)

# Summary table:

| **Topic** | **Requirement** | **Evidence** |
| --- | --- | --- |
| ML Model Architecture | "The problem addressed by the proposed architecture (e.g., minimization, maximization, clustering, regression, classification, anomaly detection, optimal control) should be described.", | "Section 6.4.3.6 a" |
|  | "ML Model inputs (list, types, definition, units, accuracy, and resolution) should be described.", | "Section 6.4.3.6 a" |
|  | "ML Model output (list, types, definition, units, accuracy, and resolution) should be described.", | "Section 6.4.3.6 a" |
|  | "The logical structure of the proposed ML Model including the ML Model elements, their learning paradigm (e.g., supervised learning), the associated ML technology (e.g., neural networks), the function they perform in the architecture and their interfaces in terms of data flows and control flows should be described.", | "Section 6.4.3.6 a" |
|  | "Sequence constraints to be met (e.g., sequence diagram) should be described.", | "Section 6.4.3.6 a" |
| ML Model Hyperparameters | "The description of the ML Model hyperparameters should be provided.", | "Section 6.4.3.6 b" |
|  | "For example, in the case of a neural network, the description of the ML Model hyperparameters should include, but is not limited to: the input layer, the number of hidden layers with the number of neurons per layer, the output layer with its number of neurons, the type of all activation functions, weight initialization, etc.", | "Section 6.4.3.6 b" |
| ML Model Parameters | "The description of the ML Model parameters should be provided.", | "Section 6.4.3.6 c" |
|  | "For example, the parameters to be described include the weights in a neural network, the support vectors in a support vector machine, the coefficients in a linear regression or logistic regression.", | "Section 6.4.3.6 c" |
| ML Model Syntax and Semantics | "The analytical/algorithmic syntax and semantics of the ML Model, including all ML Model internal operation that are necessary to compute the output(s) of the ML Model from its inputs, are described in an unambiguous manner in the ML Model description to facilitate their implementation.", | "Section 6.4.3.6 d" |
|  | "For example, in case of a convolutional neural network the analytical/algorithmic syntax and semantics description includes, but is limited to, the ML Model parameters, average pooling, padding, stride and order of summing.", | "Section 6.4.3.6 d" |
| ML Model Replication Criterion | "The replication criterion (either exact or approximated) is defined from the ML Constituent requirements and if applicable from the ML Model requirements.", | "Section 6.4.3.6 e" |
|  | "Exact replication: In this first case, the ML Model description should contain sufficient details on the ML Model semantic to fully preserve this semantic in the implemented ML Model. For example, an exact replication criterion may be the direct and faithful implementation of the ML Model description so that the implemented ML Model meets the same performance, generalization, stability, and robustness requirements.", | "Section 6.4.3.6 e" |
|  | "Approximated replication: In this second case, the ML Model description should contain sufficient details on the ML Model semantics to approximate this semantic in the implemented ML Model with a specified tolerance. For example, an approximation metric may be expressed for a given dataset by the maximal gap between the trained ML Model outputs and the implemented ML Model outputs. The corresponding approximation replication requirement may be that this maximal gap should not exceed a given value epsilon.", | "Section 6.4.3.6 e" |
| ML Model Execution Environment | "The execution environment of the ML Model is specified.", | "Section 6.4.3.6 f" |
| ML Model Description Content | "The ML Model description is developed by avoiding unnecessary adherence to the learning environment and specific format. Any necessary dependence to the learning environment (e.g., library, format) is explicitly mentioned. For example, the ML Model description may be expressed in a custom Json file with clear semantics.", | "Section 6.4.3.6 g" |
| ML Model Description Content | "Any information that should not be part of the implemented ML Model is removed or explicitly identified as \"not part of the ML Model description\". Common examples are comments, illustration, informative metrics, tool traces/logs, etc.", | "Section 6.4.3.6 h" |