**3GPP TSG SA WG5 Meeting #156 S5-244878**

**Maastricht, The Netherlands 19 - 23 August 2024 Revision of S5-244161**

|  |
| --- |
| *CR-Form-v12.1* |
| **CHANGE REQUEST** |
|  |
|  |  | **CR** |  **0161** | **rev** | **1** | **Current version:** | **18.4.0** |  |
|  |
| *For* [***HE******LP***](http://www.3gpp.org/3G_Specs/CRs.htm#_blank)*on using this form: comprehensive instructions can be found at* [*http://www.3gpp.org/Change-Requests*](http://www.3gpp.org/Change-Requests)*.* |
|  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Proposed change affects:*** | UICC apps |  | ME |  | Radio Access Network | **x** | Core Network | **x** |

|  |
| --- |
|  |
| ***Title:***  | Rel-18 CR TS28.105 corrections to ML model lifecycle figure and corresponding description |
|  |  |
| ***Source to WG:*** | Ericsson, NEC, Intel |
| ***Source to TSG:*** | S5 |
|  |  |
| ***Work item code:*** |  |  | ***Date:*** | 2024-08-05 |
|  |  |  |  |  |
| ***Category:*** |  |  | ***Release:*** | Rel-18 |
|  | *Use one of the following categories:****F*** *(correction)****A*** *(mirror corresponding to a change in an earlier release)****B*** *(addition of feature),* ***C*** *(functional modification of feature)****D*** *(editorial modification)*Detailed explanations of the above categories canbe found in 3GPP [TR 21.900](http://www.3gpp.org/ftp/Specs/html-info/21900.htm). | *Use one of the following releases:Rel-8 (Release 8)Rel-9 (Release 9)Rel-10 (Release 10)Rel-11 (Release 11)…Rel-15 (Release 15)Rel-16 (Release 16)Rel-17 (Release 17)Rel-18 (Release 18)* |
|  |  |
| ***Reason for change:*** | Further corrections to align the description with figure for ML LCM. Description for some of the seuence in the figure are missing. Direction of the sequence between ML model training and AI/ML inference is incorrect.  |
|  |  |
| ***Summary of change:*** | * Added missing description for some of the flow seuence.
* Correct a typo in Figure 4a.0-1: ML model lifecycle for the sequence between ML model training and AI/ML emulation.
 |
|  |  |
| ***Consequences if not approved:*** | Confusion and misunderstanding of ML model LCM. |
|  |  |
| ***Clauses affected:*** | 4a.0 |
|  |  |
|  | **Y** | **N** |  |  |
| ***Other specs*** |  |  |  Other core specifications  | TS/TR ... CR ...  |
| ***affected:*** |  |  |  Test specifications | TS/TR ... CR ...  |
| ***(show related CRs)*** |  |  |  O&M Specifications | TS/TR ... CR ...  |
|  |  |
| ***Other comments:*** |  |
|  |  |
| ***This CR's revision history:*** |  |

***Start of change***

4a.0 ML model lifecycle

AI/ML techniques are widely used in 5GS (including 5GC, NG-RAN, and management system), the generic AI/ML operational workflow in the lifecycle of an ML model, is depicted in Figure 4a.0-1.



**Figure 4a.0-1: ML model lifecycle**

The ML model lifecycle includes training, testing, emulation, deployment, and inference. These steps are briefly described below:

**- ML model training:** training, including initial training and re-training, of an ML model or a group of ML models. It also includes validation of the ML model to evaluate the performance when the ML model performs on the training data and validation data. If the validation result does not meet the expectation (e.g., the variance is not acceptable), the ML model needs to be re-trained.

**- ML model testing:** testing of a validated ML model to evaluate the performance of the trained ML model when it performs on testing data. If the testing result meets the expectations, the ML model may proceed to the next step If the testing result does not meet the expectations, the ML model needs to be re-trained.

**- AI/ML inference emulation:** running an ML model for inference in an emulation environment. The purpose is to evaluate the inference performance of the ML model in the emulation environment prior to applying it to the target network or system. If the emulation result does not meet the expectation (e.g., inference performance does not meet the target, or the ML model negatively impacts the performance of other existing functionalities) the ML model needs to be re-trained.

NOTE: The AI/ML inference emulation is considered optional and can be skipped in the AI/ML operational workflow.

**- ML model deployment:** ML model deployment includes the ML model loading process (a.k.a. a sequence of atomic actions) to make a trained ML model available for use at the target AI/ML inference function.

ML model deployment may not be needed in some cases, for example when the training function and inference function are co-located.

**- AI/ML inference:** performing inference using a trained ML model by the AI/ML inference function. The AI/ML inference may also trigger model re-training or update based on e.g., performance monitoring and evaluation.

NOTE: depending on system implementation and AI/ML functionality arrangments, both AI/ML inference emulation and ML deployment steps may be skiped.

***End of changes***