**3GPP TSG-SA5 Meeting #156 *S5-244619***

**Maastricht, The Netherlands, 19 - 23 August 2024** revision of S5-243095

**Source: CATT, Ericsson**

**Title: Rel-19 pCR TR 28.858 Add use case on ML model distributed training**

**Document for: Approval**

**Agenda Item: 6.19.1**

# 1 Decision/action requested

***Approval***

# 2 References

[1] Draft TR 28.858 v0.1.0 ([S5-243522](https://www.3gpp.org/ftp/TSG_SA/WG5_TM/TSGS5_156/Docs/S5-243522.zip)) : Study on AI/ML management phase 2

[2] SP-231780: "New SID: Study on AI/ML management - phase 2 "

[3] S5-243095-pCR TR 28.908 Add use case on ML model distributed training

# 3 Rationale

It is proposed to add a new use case on ML entity distributed training.

# 4 Detailed proposal

This contribution proposes to make the following changes in [1].

|  |
| --- |
| **1st change** |

### 5.1.x ML model distributed training

#### 5.1.x.1 Description

Distributed training is a model training paradigm that involves spreading training workload across multiple training functions, to accelerate the training process and/or reduce the required computational resources. Distributed training can be used for traditional machine learning models, as well as for large models.

In 5GS, the ML training function may be located within the management system or in the NF (e.g. gNB or NWDAF), i.e. the worker node for training. Each node has different computing resources and storage capacity based on physical infrastructure such as CPU/GPU/DPU, memory, storage, and network bandwidth. In order to obtain load balance between nodes and maximize the efficiency of resource utilization, splitting up the training may be necessary and involving multiple training functions according to the actual situation of nodes may be needed. Thus, aspects of distributed training need to be supported in the management systems.

#### 5.1.x.2 Use cases

#### 5.1.x.2.1 ML model distributed training

In 5GS, the MnS consumer may require to speed up the training process under certain conditions (e.g. the size of the model may be too large for a single training function).

When receiving an ML training request, the MLT MnS producer may evaluate whether distributed training is needed according to the training requirements provided by the ML training consumer, and it is up to the MLT MnS producer to determine appropriate training function(s) which need to participate in the ML model training. Collaboration and mutual agreement may require between distributed ML training functions.

NOTE 1: How to split the ML model and synchronize the parameters in different training function depends on the distributed algorithm which are proprietary and not in scope for standardization.

NOTE 2: The data exchange between different training functions should be in the security tunnel with appropriate authentication and authorization mechanisms.

#### 5.1.x.3 Potential requirements

**REQ-ML\_DIST-TRNG-01** The ML training MnS producer should have a capability allowing the authorized consumer to provide distributed training requirements to the MnS Producer.

#### 5.1.x.4 Possible solutions

TBD

#### 5.1.x.5 Evaluation

TBD

|  |
| --- |
| **End of change** |