3GPP TSG SA WG5 Meeting 136-e S5-212269

**electronic meeting, online, 1st – 9th March 2021**

**Source: Nokia**

**Title: ML Model training report**

**Document for: Approval**

**Agenda Item: 6.5.4**

# 1 Decision/action requested

***This contribution is for approval.***

# 2 References

[1] 3GPP TS 28.809 Management and orchestration; Study on enhancement of Management Data Analytics (MDA)

# 3 Rationale

Currently, the training result of the ML model training report produced by the ML training model towards the MDAS consumer is suggested to indicate if the training process was successful, partically sucessful or failed. This operation is invalid since a ML training process should be successful before the ML model is deployed and used. Maybe the ML training does not help to provide accurate results but that does not mean that it was not successful. In our view the ML model training report should contain the expected accuracy instead in the training result. Also, the cause of failure is difficult to capture and not sure if it is meaningful to send to the consumer but instead send the metric that is going to be used o evaluate and validate the report is more meaningful.

# 4 Detailed proposal

It is proposed to update the following chapters in TR 28.809 [1].

|  |
| --- |
| **1st Modified Section** |

### 6.99.1 ML model training for MDA

#### 6.99.1.1 Use case

The MDA process may rely on ML technologies. To optimize the accuracy of MDA result, the ML model of the MDA process may need to be trained.

For training the ML model of the MDA process, the consumer may provide the training data including training input and the desired output to the MDAS producer. The MDAS producer uses the training input and the desired output to train the ML model, i.e., to train the algorithm of the ML model to generate the desired output as accurately as possible by analysis of the training input. The MDAS producer provides an ML model training report to the consumer.

With a trained ML model for MDA, the MDAS producer can analyse the analytics input and generate the analytics report as output data of the analysis to the consumer.

The consumer may validate the output data provided by the MDAS producer. The output data to be validated may be the analytics report and/or the ML model training report as described above. The consumer may provide the validation data as feedback to the MDAS producer, and the MDAS producer will use the validation data for further ML model training for MDA with the input data that were used to generate the validated report. As a result of validation, the consumer may also provide the training data and request the MDAS producer to train the ML model.

The consumer may validate the output data provided by the MDAS producer. The output data to be validated may be the analytics report and/or the ML model training report as described above. The consumer may provide the validation data as feedback to the MDAS producer, and the MDAS producer will use the validation data for further ML model training for MDA with the input data that were used to generate the validated report. As a result of validation, the consumer may also provide the training data and request the MDAS producer to train the ML model.

#### 6.99.1.2 Potential requirements

**REQ-MDA\_MGMT-CON-1** The MDAS producer should have a capability allowing the consumer to provide the data for training the ML model for MDA.

**REQ-MDA\_MGMT-CON-2** The MDAS producer should have a capability to provide ML model training report to the consumer.

**REQ-MDA\_MGMT-CON-3** The MDAS producer should have a capability to receive the validation data from the consumer and train the ML model for MDA based on the received validation data

#### 6.99.1.3 Possible solutions

##### 6.99.1.3.1 Solution description

The MDAS producer trains the ML model with the training data or the validation data received from consumer. The ML model training should have no or minimal impact to the normal data analytics.

For the ML model training with the training data, the training data (see subclause 6.99.1.3.2) should include the training Id, the training input and expected training output:

- The training Id is used to identify the ML model training request, and to associate with the training report.

- The training input is a set of training input data with the indicated data type (e.g., performance measurements, MDT report, NRM, etc.).

- The expected training output specifies the analysis result that the ML model should aim to achieve based on the training input.

For the ML model training with the validation data, the validation data (see subclause 6.99.1.3.2) should include the validation report Id, analytics report Id that was validated, and the validated information including the data that are rectified. The MDAS producer looks up the historical data that are associated with the validated analytics report and trains the ML model with the historical data and the validation data.

The MDAS producer provides a training report (as shown in subclause 6.99.1.3.3) to the consumer, with indication of whether the training (identified by the training Id or validation Id) is successful and possibly the failure cause if the training is not fully successful.

##### 6.99.1.3.2 Data required for ML model training for MDA

The following table describes the data required for ML model training for MDA:

|  |  |
| --- | --- |
| **Data category** | **Required data** |
| Training data | The training data include a training identifier, the training input and the expected training output.  Training Id: Identifier of the training.  The training input contains the same types of information that should be included in a normal analytics input.  The expected training output contains the same types of information that should be included in a normal analytics report. |
| Validation data | The training data include a validation report Id, analytics report Id that was validated, and the data that are rectified. Example of the rectified data is the originally reported root cause is not accurate by validation and should be updated to another. |

##### 6.99.1.3.3 ML model training report

Following table provides the potential contents of the ML model training report.

|  |  |  |
| --- | --- | --- |
| **ML model training report** | **Attribute Name** | **Description** |
| Training Id | The training Id that the training report is associated to |
| Validation report Id | The validation report Id that the training report is associated to. |
| Usage of consumer provided data for training | Indication of weather the ML model is trained based on the consumer input data, e.g. training is fully or partially based on the data provided by the consumer, or the ML model is not trained based on the consumer data. |
| Deviation from consumer provided data for training | The index indicating the deviation of the training result from the data provided by the consumer for training, in case the ML model is partially trained based on the data provided by the consumer. |
| Metric value | Assuming that the MDAS producer and MDAS consumer have agreed on using metrics for evaluating the performance of the ML model, the metric value can then provide an indication regarding the performance of the ML model, e.g. the average absolute error. |

6.99.1.4 Evaluation

The solution described in clause 6.99.1.3 requires the input s described in in clause 6.99.1.3.2, wherein

- the training data are fully aligned with the analytics input and output of the concrete analytics use cases.

- and validation data can be provided by the consumer.

Therefore, this solution is a feasible candidate for ML model training for the use cases that are concluded with a feasible solution.

|  |
| --- |
| **End of Modified Sections** |