



INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

Y.3176

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

(09/2020)

SERIES Y: GLOBAL INFORMATION
INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS,
NEXT-GENERATION NETWORKS, INTERNET OF
THINGS AND SMART CITIES

Future networks

Machine learning marketplace integration in future networks including IMT-2020

CAUTION !

PREPUBLISHED RECOMMENDATION

This prepublication is an unedited version of a recently approved Recommendation. It will be replaced by the published version after editing. Therefore, there will be differences between this prepublication and the published version.



Documents

Document Sharing Hub

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU [had/had not] received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at <http://www.itu.int/ITU-T/ipr/>.

© ITU 2020

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Draft new Recommendation ITU-T Y.3176 (formerly Y.ML-IMT2020-MP)

Machine learning marketplace integration in future networks including IMT-2020

Summary

This Recommendation provides high-level requirements and the architecture for integrating ML marketplaces in future networks including IMT-2020. Based on these requirements, the architecture for the integration of ML marketplaces is described taking into account the architectural framework in [ITU-T Y.3172] as a basis.

Keywords

architecture, federation, future networks, IMT-2020, life-cycle, machine learning, management, marketplace, model, network functions, requirements

Draft new Recommendation ITU-T Y.3176 (formerly Y.ML-IMT2020-MP)

Machine learning marketplace integration in future networks including IMT-2020

1 Scope

This Recommendation provides high-level requirements and the architecture for integration of ML marketplaces in future networks including IMT-2020.

The scope of this Recommendation includes:

- Main motivations and challenges related to ML marketplace integration in networks;
- High level requirements of ML marketplace integration;
- Architecture for the integration of ML marketplaces in networks.

2 References

The following ITU-T Recommendations and other references contain provisions, which, through reference in this text, constitute provisions of this deliverable. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this deliverable are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published.

The reference to a document within this deliverable does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T Y.2701]	ITU-T Recommendation Y.2701 (2007), “ <i>Security requirements for NGN release 1</i> ”
[ITU-T Y.3100]	ITU-T Recommendation Y.3100 (2017), “ <i>Terms and definitions for IMT-2020 network</i> ”
[ITU-T Y.3101]	ITU-T Recommendation Y.3101 (2018), “ <i>Requirements of the IMT-2020 network</i> ”
[ITU-T Y.3104]	ITU-T Recommendation Y.3104 (2018), “ <i>Architecture of the IMT-2020 network</i> ”
[ITU-T Y.3110]	ITU-T Recommendation Y.3110 (2017), “ <i>IMT-2020 network management and orchestration requirements</i> ”
[ITU-T Y.3111]	ITU-T Recommendation Y.3111 (2017), “ <i>IMT-2020 network management and orchestration framework</i> ”
[ITU-T Y.3172]	ITU-T Recommendation Y.3172 (2019), “ <i>Architectural framework for machine learning in future networks including IMT-2020</i> ”
[ITU-T Y.3174]	ITU-T Recommendation Y.3174 (2020), “ <i>Framework for data handling to enable machine learning in future networks including IMT-2020</i> ”

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 machine learning (ML) [ITU-T Y.3172]: processes that enable computational systems to understand data and gain knowledge from it without necessarily being explicitly programmed.

NOTE 1 – Definition adapted from [b-ETSI GR ENI 004].

NOTE 2 – Supervised machine learning and unsupervised machine learning are two examples of machine learning types.

3.1.2 machine learning database [ITU-T Y.3174]: A component which stores data related to machine learning in the machine learning overlay.

3.1.3 machine learning function orchestrator (MLFO) [ITU-T Y.3172]: a logical node with functionalities that manage and orchestrate the nodes in a machine learning pipeline.

3.1.4 machine learning model [ITU-T Y.3172]: model created by applying machine learning techniques to data to learn from.

NOTE 1 – A machine learning model is used to generate predictions (e.g. regression, classification, clustering) on new (untrained) data.

NOTE 2 – A machine learning model may be encapsulated in a deployable fashion in the form of a software (e.g. virtual machine, container) or hardware component (e.g. IoT device).

NOTE 3 – Machine learning techniques include learning algorithms (e.g. learning the function that maps input data attributes to output data).

3.1.5 machine learning pipeline [ITU-T Y.3172]: A set of logical nodes, each with specific functionalities, that can be combined to form a machine learning application in a telecommunication network.

NOTE – The nodes of a machine learning pipeline are entities that are managed in a standard manner and can be hosted in a variety of network functions [ITU-T Y.3100].

3.1.6 machine learning sandbox [ITU-T Y.3172]: an environment in which machine learning models can be trained, tested and their effects on the network evaluated.

NOTE – A machine learning sandbox is designed to prevent a machine learning application from affecting the network, or to restrict the usage of certain machine learning functionalities.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 machine learning model metadata: information which describes the characteristics of a machine learning model.

NOTE – Machine learning (ML) model metadata include, but is not limited to, name of the ML model, ML model's author, version of the ML model, license information of the ML model, description of the data inputs and outputs of the ML model and runtime environment of the ML model.

3.2.2 machine learning marketplace: a component which provides capabilities facilitating the exchange and delivery of machine learning models among multiple parties.

Rec. ITU-T Y.3176 (09/2020) – Prepublished version

NOTE 1 – Examples of parties include suppliers and users of ML models. Capabilities provided to users of ML models include functionalities to find, learn about, deploy (or download) and use ML models. Capabilities provided to suppliers of ML models (e.g. data scientist) include functionalities to share (on-board, upload), describe (learn about) and market their ML models.

NOTE 2 – A network operator may use a machine learning marketplace deployed internally and/or externally to the network operator's administrative domains. Internal and external marketplaces differ only in the deployment perspective. A marketplace which is internal to a network operator may act as an external marketplace to another network operator and vice versa.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

API	Application Programming Interface
CN	Core Network
Enum	Enumeration
ID	Identifier
KPI	Key Performance Indicator
ML	Machine Learning
MLFO	Machine Learning Function Orchestrator
MLDB	Machine Learning Database
MPP	Mobility Pattern Prediction
SRC	Source
TCP	Transmission Control Protocol
URI	Uniform Resource Identifier

5 Conventions

In this Recommendation, requirements are classified as follows:

- The keywords "**is required to**" indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this Recommendation is to be claimed.
- The keywords "**is recommended**" indicate a requirement which is recommended but which is not absolutely required. Thus, such requirements need not be present to claim conformance.
- The keywords "can optionally" and "**may**" indicate an optional requirement which is permissible, without implying any sense of being recommended. These terms are not intended to imply that the vendor's implementation must provide the option and the feature can be optionally enabled by the NOP/service provider. Rather, it means the vendor may optionally provide the feature and still claim conformance with the specification.

6 Introduction

Use cases for use of machine learning (ML) in future networks including IMT-2020 are described in [b-ITU-T Sup.55] while an architectural framework for ML in future networks including IMT-2020 is presented in [ITU-T Y.3172].

The main motivations for ML marketplace integration in future networks including IMT-2020 are highlighted below:

1. Evolution of ML models may happen independently of the technology evolution in future ML underlay networks. Integration of ML marketplaces will provide a mechanism for network operators to follow the innovation curve in the domain of machine learning.

NOTE – For example, as and when new, innovative and proven ML mechanisms and techniques are published by ML researchers and corresponding ML models are made available in the marketplace, network operators are able to use those models using ML marketplace integration described in this Recommendation.

2. Efficient exchange and deployment of ML models using standard interfaces will help not only solving network problems using ML techniques but also in sharing and monetizing ML techniques in the network domain.

However, the following main challenges need to be addressed for the integration of ML marketplaces in ML-enabled networks:

- ML models may be hosted in varied types of ML marketplaces (e.g. Acumos [b-LF Acumos]). While designing the ML application, network operators need interoperable mechanisms for identification of the ML marketplace which may be used as a source for ML models.

NOTE 1 – This points to the need for network operators to use standard ML intent mechanisms to specify the ML marketplace to be used for the ML application.

- Network operators need interoperable, automated mechanisms to select the appropriate ML models from the ML marketplaces.

NOTE 2 – For example, the latest advances in analytics or algorithms [b-DL Survey] have no dependency on the network architecture evolution of ML underlay networks. Cloud computing-hosted ML marketplaces may attract developers of innovative ML mechanisms and algorithms to host their ML models.

NOTE 3 – Metadata used for describing ML models stored in ML marketplaces may be exploited for the selection process. In such a context, availability of standardized ML model metadata is of prime importance.

- Network operators should be able to follow the innovation curve in the ML marketplace without sacrificing their network key performance indicators (KPI) [b-3GPP TS 28.554].

NOTE 4 – Network operators may use a ML sandbox to isolate the effect of training and testing of ML models from the live network.

- Lack of standard mechanisms to exchange ML models and related metadata between ML marketplaces and the network operator's ML deployment environments limits the interoperability. There is a need to study and develop standard mechanisms for the interfaces between ML marketplaces and network operator's ML deployment environments in order to allow for efficient exchange and deployment of ML models.

NOTE 5 – For example, exchange and deployment of ML models from a ML marketplace may be done using export of ML models to standard file format [b-ONNX], ML model

compression using standard methods [b-ISO-IEC 15938-17] and secure file transfer to network operator's ML deployment environments.

NOTE 6 – Deployment of internal ML marketplaces enables collaborative development and use of ML models within different ML deployment environments of the network operator, e.g. operations and maintenance, research and development. It also enables the development of regional or enterprise specific ML models which conform to local regulations and organizational policies on data handling.

- As mentioned in [ITU-T Y.3172], building complex ML models by chaining ML models is an important need for network operators. But selecting the appropriate ML models from disparate ML marketplaces and chaining them requires support for such mechanisms from ML marketplaces.

NOTE 7 – This points to the need for the availability of standardized metadata for the ML models and the need to interoperate with ML sandbox.

- Efficient use of testing methods requires sharing of information between ML marketplaces and network operators, e.g. data which was used for testing, test results, current KPIs for the ML model.
- Deployment of ML models depends on the targeted deployment environment of the network operator, which differ from a virtualisation technology standpoint, e.g. container or virtual machine (VM) based [b-ETSI-GS-NFV-EVE]. ML models hosted in ML marketplaces need to be amenable to such different types of deployment environment.

NOTE 8 – For example, a containerization of ML model may be needed to deploy the model in a container-based virtualization environment of the network operator.

This Recommendation provides high-level requirements and architectural components for integrating ML marketplaces in future networks including IMT-2020.

7 High-level requirements for ML marketplace integration

This clause provides high-level requirements for ML marketplace integration, i.e. ML model metadata related requirements and ML marketplace requirements.

7.1 ML model metadata related requirements

REQ-ML-MKT-METADATA-01: The ML model metadata is required to indicate the training status and chaining status of the ML model.

NOTE 1 – Examples of indicated training status information include whether the ML model is trained or not, and in case of trained models, the training datasets used.

NOTE 2 – Examples of indicated chaining status information include whether the ML model is compatible with any other model to be chained.

REQ-ML-MKT-METADATA-02: The ML model metadata is required to indicate the testing maturity of the model.

NOTE 3 – Examples of testing maturity indications of a ML model include the following:

- 1) Performance benchmarks with respect to published or standard test cases, e.g. latency benchmarks of the ML model, memory requirements of the ML model, accuracy of the ML model.
- 2) Reference to standard test cases and corresponding compliance of the ML model to those test cases.

3) Applicability of ML output from the ML model to various types of ML underlay networks, e.g. ML output from mobility pattern prediction (MPP) model may be applicable to urban and rural areas.

4) A rank or score based on testing or user experience from network operators.

NOTE 4 – For example, a network operator may score the ML model (based on the use case under consideration) as “10” on a scale from “0” (lowest score) to “10” (top score). This score may be used by another network operator to select the ML model (to be used for a similar use case) from the ML marketplace.

REQ-ML-MKT-METADATA-03: The ML model metadata is required to indicate the input and output characteristics of the ML model.

NOTE 5 – Examples of indicated characteristics include the dimensions of input data and ML output, and data models used in exchange of data with ML underlay [ITU-T Y.3174].

REQ-ML-MKT-METADATA-04: The ML model metadata is required to indicate the compression status of the ML model.

NOTE 6 – Examples of indicated compression status information include whether the ML model is compressed or not, and in case of compressed models, the compression method used.

REQ-ML-MKT-METADATA-05: The ML model metadata is required to indicate the type of learning that the ML model supports.

NOTE 7 – Examples of learning type indication include supervised or unsupervised learning.

NOTE 8 – The indication about the learning type is used for the training or retraining of the ML model. Depending on the type of learning indicated, the MLFO is able to select an appropriate type of ML pipeline to be instantiated in order to train such an ML model.

REQ-ML-MKT-METADATA-6: The ML model metadata is required to indicate the resource requirements and constraints for ML models.

NOTE 9 – Specified resource requirements and constraints include compute, storage and network requirements for the ML model to function and to meet the KPIs of ML use cases. Examples are deployment environment of the ML model, prediction time, budget for communication latency, and validity time for ML output. Example of temporal validity for ML output is the case where prediction from the ML model may be valid for use only for a limited time, after that the prediction needs to be refreshed based on new data.

7.2 ML marketplace related requirements

REQ-ML-MKTPLACE-01: The ML marketplace is recommended to support functionalities for the searching of ML models based on the metadata describing these ML models.

NOTE 1 – For example, a ML marketplace may support an interface that takes metadata (or its components) as input for selection and returns a list of ML models satisfying the best match for those metadata.

REQ-ML-MKTPLACE-02: The ML marketplace is recommended to support functionalities for exporting ML models according to standardized formats.

NOTE 2 – Example of standardized model format for exporting includes ONNX [b-ONNX].

REQ-ML-MKTPLACE-03: The ML marketplace is recommended to support asynchronous notifications to the MLFO about additions, modifications and deletions of new ML models.