

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Y.3172

(06/2019)

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

Future networks

Architectural framework for machine learning in future networks including IMT-2020

Recommendation ITU-T Y.3172

ITU-T Y-SERIES RECOMMENDATIONS

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

GLOBAL INFORMATION INFRASTRUCTURE	
General	Y.100–Y.199
Services, applications and middleware	Y.200–Y.299
Network aspects	Y.300–Y.399
Interfaces and protocols	Y.400–Y.499
Numbering, addressing and naming	Y.500–Y.599
Operation, administration and maintenance	Y.600–Y.699
Security	Y.700–Y.799
Performances	Y.800–Y.899
INTERNET PROTOCOL ASPECTS	
General	Y.1000–Y.1099
Services and applications	Y.1100–Y.1199
Architecture, access, network capabilities and resource management	Y.1200–Y.1299
Transport	Y.1300–Y.1399
Interworking	Y.1400–Y.1499
Quality of service and network performance	Y.1500–Y.1599
Signalling	Y.1600–Y.1699
Operation, administration and maintenance	Y.1700–Y.1799
Charging	Y.1800–Y.1899
IPTV over NGN	Y.1900–Y.1999
NEXT GENERATION NETWORKS	
Frameworks and functional architecture models	Y.2000–Y.2099
Quality of Service and performance	Y.2100–Y.2199
Service aspects: Service capabilities and service architecture	Y.2200–Y.2249
Service aspects: Interoperability of services and networks in NGN	Y.2250–Y.2299
Enhancements to NGN	Y.2300–Y.2399
Network management	Y.2400–Y.2499
Network control architectures and protocols	Y.2500–Y.2599
Packet-based Networks	Y.2600–Y.2699
Security	Y.2700–Y.2799
Generalized mobility	Y.2800–Y.2899
Carrier grade open environment	Y.2900–Y.2999
FUTURE NETWORKS	Y.3000–Y.3499
CLOUD COMPUTING	Y.3500–Y.3999
INTERNET OF THINGS AND SMART CITIES AND COMMUNITIES	
General	Y.4000–Y.4049
Definitions and terminologies	Y.4050–Y.4099
Requirements and use cases	Y.4100–Y.4249
Infrastructure, connectivity and networks	Y.4250–Y.4399
Frameworks, architectures and protocols	Y.4400–Y.4549
Services, applications, computation and data processing	Y.4550–Y.4699
Management, control and performance	Y.4700–Y.4799
Identification and security	Y.4800–Y.4899
Evaluation and assessment	Y.4900–Y.4999

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T Y.3172

Architectural framework for machine learning in future networks including IMT-2020

Summary

Recommendation ITU-T Y.3172 specifies an architectural framework for machine learning (ML) in future networks including IMT-2020. A set of architectural requirements and specific architectural components needed to satisfy these requirements are presented. These components include, but are not limited to, an ML pipeline as well as ML management and orchestration functionalities. The integration of such components into future networks including IMT-2020 and guidelines for applying this architectural framework in a variety of technology-specific underlying networks are also described.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T Y.3172	2019-06-22	13	11.1002/1000/13894

Keywords

Architectural framework, high-level architecture, IMT-2020, intent, machine learning, ML, orchestrator, overlay, pipeline, requirements, sandbox, training, underlay.

* To access the Recommendation, type the URL <http://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID. For example, <http://handle.itu.int/11.1002/1000/11830-en>.

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 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 2019

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

Table of Contents

		Page
1	Scope.....	1
2	References.....	1
3	Definitions	1
	3.1 Terms defined elsewhere	1
	3.2 Terms defined in this Recommendation.....	1
4	Abbreviations and acronyms	2
5	Conventions	3
6	Introduction.....	4
7	High-level architectural requirements	5
	7.1 Enablers for correlation	5
	7.2 Enablers for deployment.....	7
	7.3 Interface-related.....	9
	7.4 Declarative specifications.....	10
	7.5 Management of ML functionalities	12
8	Framework of the high-level architecture.....	13
	8.1 High-level architectural components.....	14
	8.2 High-level architecture	16
	8.3 General guidelines for realization of the high-level architecture	19
9	Security considerations	19
	Appendix I – Realization of the high-level architecture on technology-specific underlay networks.....	21
	Appendix II – Mapping of the architectural components and supporting aspects to the requirements in clause 8	23
	Bibliography.....	25

Recommendation ITU-T Y.3172

Architectural framework for machine learning in future networks including IMT-2020

1 Scope

This Recommendation specifies an architectural framework for machine learning in future networks including IMT-2020.

A set of architectural requirements is presented, which in turn leads to specific architectural components needed to satisfy these requirements. This Recommendation also describes an architectural framework for the integration of such components into future networks including IMT-2020 and guidelines for applying this architectural framework in a variety of technology-specific underlying networks.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation 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 Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T Y.3111] Recommendation ITU-T Y.3111 (2017), *IMT-2020 network management and orchestration framework*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following term defined elsewhere:

3.1.1 network function [b-ITU-T Y.3100]: In the context of IMT-2020, a processing function in a network.

NOTE 1 – Network functions include but are not limited to network node functionalities, e.g., session management, mobility management and transport functions, whose functional behaviour and interfaces are defined.

NOTE 2 – Network functions can be implemented on a dedicated hardware or as virtualized software functions.

NOTE 3 – Network functions are not regarded as resources, but rather any network functions can be instantiated using the resources.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 machine learning (ML): Processes that enable computational systems to understand data and gain knowledge from it without necessarily being explicitly programmed.

NOTE 1 – This definition is adapted from [b-ETSI GR ENI 004].

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

3.2.2 machine learning function orchestrator (MLFO): A logical node with functionalities that manage and orchestrate the nodes in a machine learning pipeline.

3.2.3 machine learning model: 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).

3.2.4 machine learning overlay: A loosely coupled deployment model of machine learning functionalities whose integration and management with network functions are standardised.

NOTE – A machine learning overlay aims to minimise interdependencies between machine learning functionalities and network functions using standard interfaces, allowing for the parallel evolution of functionalities of the two.

3.2.5 machine learning pipeline: 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 [b-ITU-T Y.3100].

3.2.6 machine learning sandbox: 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.

3.2.7 machine learning underlay network: A telecommunication network and its related network functions which interfaces with corresponding machine learning overlays.

NOTE – An IMT-2020 network is an example of a machine learning underlay network.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AF	Application Function
AN	Access Network
API	Application Programming Interface
AR/VR	Augmented Reality/Virtual Reality
C	Collector (in ML pipeline)
CN	Core Network
D	Distributor (in ML pipeline)
EMS	Element Management System
FMC	Fixed Mobile Convergence
GPU	Graphic Processor Unit
IoT	Internet of Things
KPI	Key Performance Indicator
M	Model (in ML pipeline)

mIoT	massive Internet of Things
MEC	Multi-access Edge Computing
ML	Machine Learning
MLFO	Machine Learning Function Orchestrator
MnS	Management Service
MPP	Mobility Pattern Prediction
NF	Network Function
NOP	Network Operator
OAM	Operations, Administration and Maintenance
P	Policy (in ML pipeline)
PP	Preprocessor (in ML pipeline)
QoS	Quality of Service
RCA	Root Cause Analysis
RRC	Radio Resource Control
SBA	Service-Based Architecture
SMF	Session Management Function
SON	Self-Optimizing Network
SRC	Source
UE	User Equipment
UPF	User Plane Function
V2X	Vehicle-to-everything
VoLTE	Voice over Long-Term Evolution

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, this requirement need not be present to claim conformance.

The keywords "can optionally" indicate an optional requirement which is permissible, without implying any sense of being recommended. This term is not intended to imply that the vendor's implementation must provide the option and the feature can be optionally enabled by the network operator/service provider. Rather, it means the vendor may optionally provide the feature and still claim conformance with the specification.

In the body of this Recommendation and its annexes, the words shall, shall not, should, and may sometimes appear, in which case they are to be interpreted, respectively, as is required to, is prohibited from, is recommended, and can optionally. The appearance of such phrases or keywords in an appendix or in material explicitly marked as informative is to be interpreted as having no normative intent.

ML pipeline – In this Recommendation, when the symbol shown in Figure 1 is used, this denotes a subset (including proper subset) of nodes in an ML pipeline. When this symbol is used in a figure, the symbol stands for the subset of an ML pipeline's nodes not explicitly shown in that figure.



Figure 1 – Symbol used to denote a subset of nodes in an ML pipeline

Service egress and ingress points – In this Recommendation, a node and its service egress and ingress points are denoted by the symbol in Figure 2. The service egress point is shown with a green rectangle on top of the node, while the service ingress point corresponds to the red rectangle on top of the node.

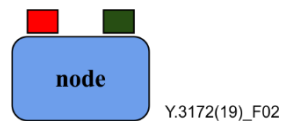


Figure 2 – Symbol used to denote a node with its service egress and ingress points

6 Introduction

Machine learning (ML) provides a way to teach computational systems to gain knowledge from data without necessarily being explicitly programmed, in order to realize complicated tasks such as detection of characteristics or prediction of behaviours. As ML becomes an important technical trend in the industry, network operators and other stakeholders are searching for cost-effective ways to incorporate ML into future networks including IMT-2020.

While the benefits from such an integration have been discussed under many use cases (e.g., troubleshooting of network problems, network traffic prediction, traffic optimization adjustment, network security auditing [b-ITU-T Y.3650] [b-ITU-T Y.3170]), there are many challenges to such an integration. Some of the important challenges are:

- A. The heterogeneous nature of ML functionalities and unique characteristics of future communication technologies impose a varied set of requirements for integration.
- B. Roadmaps for evolution of these ML functionalities and communication networks are not aligned.
- C. The cost of integration, in terms of architecture impact, is an important consideration.
- D. Disparate management mechanisms for ML functionalities and network functions [b-ITU-T Y.3100] will disrupt the operations management of communication networks.

An architectural framework for the integration of ML with future networks including IMT-2020 is provided to address these challenges.

Building on high-level architectural requirements, such a framework at first provides a common vocabulary and nomenclature for ML functionalities and their relationships with the communication networks. The framework addresses the reference points for future networks including IMT-2020 which enable loosely coupled integration with ML functionalities. Management mechanisms for ML are also described in the framework considering the existing management principles which have been identified in IMT-2020 networks.

The framework enables a standard method of integrating ML functionalities in future networks including IMT-2020.

7 High-level architectural requirements

This clause provides high-level requirements for the design of the high-level architecture framework in clause 8.

These high-level architectural requirements are classified as shown below.

- Enablers for correlation of data across levels and heterogeneous technologies: This set of requirements addresses challenge A mentioned in clause 6 and the requirements which fall into this set are provided in clause 7.1.

NOTE – Levels may correspond to different parts of the network where the reference points between these parts are defined in the relevant network architecture specifications. For example, [b-ITU-T Y.3104] defines the reference point RP-an between the access network (AN) and the core network (CN), thus, AN and CN may correspond to different levels of an IMT-2020 network.

- Enablers for deployment: This set of requirements addresses challenge B mentioned in clause 6 and the requirements which fall into this set are provided in clause 7.2.
- Requirements related to interfaces between the architectural components: This set of requirements addresses challenges B and C mentioned in clause 6 and the requirements which fall into this set are provided in clause 7.3.
- Requirements related to declarative specifications used for specifying the ML applications: This set of requirements addresses challenges C and D mentioned in clause 6 and the requirements which fall into this set are provided in clause 7.4.
- Requirements related to the management of the architectural components: This set of requirements addresses challenge D mentioned in clause 6 and the requirements which fall into this set are provided in clause 7.5.

7.1 Enablers for correlation

Table 7-1 provides requirements regarding enablers for the correlation of data across levels and heterogeneous technologies.

Table 7-1 – High-level requirements – Enablers for correlation

REQ-ML-COR-001	The ML architecture is recommended to support the correlation of data coming from multiple sources.
Description	In future networks, sources of data may be heterogeneous, integrated with different network functions (NFs), and may report different formats of data. These varied "perspectives" can provide rich insights upon correlated analysis. Architectural components to enable the ML functionalities to collect and correlate data from these varied sources in the network are needed.
Notes	NOTE 1 – As an example, the analysis of data from user equipment (UE), AN, CN and application function (AF) is needed to predict potential issues related to quality of service (QoS) in end-to-end user flows. NOTE 2 – Other examples of such sources of data are self-optimizing network (SON) functionalities that monitor and correlate network alarms and key performance indicators (KPIs) [b-3GPP TS 28.554], and then take relevant action to clear alarms or enhance network KPIs, or give network design recommendations without human intervention.