SA WG2 Meeting #163 S2-2406912

Jeju island, Korea, May 27 to May 31, 2024 (Revision of S2-2406316)

**Source: China Mobile, ZTE, OPPO, CATT, vivo, Huawei, Sony**

**Title: VFL terminology definition**

**Document for: Approval**

**Agenda Item: 19.15**

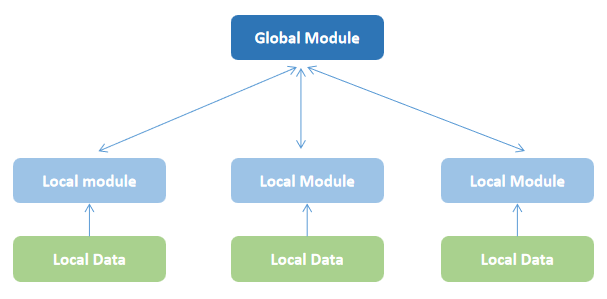
**Work Item / Release: FS\_AIML\_CN/Rel-19**

*Abstract of the contribution: This contribution* *proposes the VFL terminology definition.*

# 1. Introduction/Discussion

**1.1 VFL process**

Vertical Federated Learning (VFL) training process is illustrated as follows:



The model consists of a global module and multiple local modules. The local modules are trained using local datasets at first, and the intermediate results are sent to the global module for aggregation.

**1.2 The understanding of VFL Active Participant/VFL Passive Participant, VFL Server/VFL Client, and Coordinator**

We distinguish the terms in the view of model:

VFL Active Participant is the entity which can hold both the label and optionally one of the local modules. It collects local data and also gets intermediate results from the other local modules. And VFL Passive Participant is the entity that holds the local module. (The definition of label is not clear currently. It is proposed to change it as training objective in supervised machine learning.)

VFL Server is the entity that performs as a VFL Active Participant, holding both the global module and optionally one of the local modules. Apart from the computing functions, control functions are also needed in VFL Server to help the VFL system work, which we called as a coordinator. It coordinates the VFL process and takes control of all the VFL participants.

And VFL Client is the same as VFL Passive Participant, only holding the local module.

**1.3 Conclusion**

In conclusion, VFL Server is a more general term, because it can perform as a VFL Active Participant and also coordinator. In this way, VFL Server plays the role of both the computing and controlling. Therefore, we propose to use the term of VFL Server/VFL Client in the TR.

Using the term of VFL Server/VFL Client makes it easier to align VFL with HFL. For example, a centralized node can served both as a HFL Server and a VFL Server without major changes, due to that the service operations defined for HFL can be reproduced. It would be better to work with the existing Federated Learning framework than introducing new confusing terms (e.g., Active/Passive Participant) which may disrupt the current system.

# Text Proposal

It is proposed to agree the following changes to TR 23.700-84.

\* \* \* \* First change \* \* \* \*

## 3.1 Terms

For the purposes of the present document, the terms given in TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in TR 21.905 [1].

**Horizontal Federated Learning (HFL):** a federated learning technique without exchanging/sharing local data set, wherein the local data set in different FL clients for local model training have the same feature space for different samples (e.g. UE IDs).

**Vertical Federated Learning (VFL):** a federated learning technique without exchanging/sharing local data set, wherein the local data set in different VFL Participant for local model training have different feature spaces for the same samples (e.g. UE IDs).

**Label:** A label is the training objective in supervised machine learning.

**VFL Active Participant:** An NF with labels for a VFL training task that may have related input data.

**VFL Passive Participant:** A VFL client with access to the required input data without the required labels for a VFL training task. There can be multiple passive participants in VFL.

NOTE 1: It will be determined in the normative phase whether definitions of VFL Active Participant and VFL passive participant are required in normative specs.

VFL Server: In VFL training process, the VFL server is an NF, i.e., NWDAF or AF, integrating all the local training results and computing gradient information or loss information for the local ML model update. It also coordinates the VFL training process by discovering and selecting VFL clients. In VFL inference process, The VFL server aggregates local inference results from VFL clients to generate the final VFL inference result and sends the final VFL inference result to the VFL inference consumer.

NOTE 2: It will be determined in the normative phase whether a separation into VFL training server and VFL inference server is required.

VFL Client: An NF, i.e., NWDAF or AF, which holds the local dataset and performs local training and inference. There can be multiple VFL Clients in VFL training and inference. The VFL client does not have label.

NOTE 3: It will be determined in the normative phase whether a separation into VFL training client and VFL inference client is required.

\* \* \* \* End of changes \* \* \* \*