Source: https://challenge.aiforgood.itu.int/match/matchitem/84

With the wide spread of Network Function Virtualization (NFV) technology, a plenty of Virtualized Network Functions (VNFs) and Cloud-native Network Functions (CNFs) has been deployed in telco 5G mobile networks. In future, a greater number of VNFs/CNFs are expected to be integrated in the Beyond 5G (B5G) networks. This trend is causing the increase in operational workloads due to the complexity and dynamic changes in networks. Network Digital Twin (NDT), or digital twin network [1], is expected to support these operational tasks such as failure analysis and mitigation to maintain high-level requirements for B5G. NDT is a mirror network for a real network, which realizes a variety of practical simulations related to network behaviours and operations such as upgrading and disaster recovery to help actual operations in a real environment. One approach using NDT is to develop AI/ML models in NDT to apply in a real network. Since it is difficult to prepare the completely same network as well as to simulate all the possible patterns in NDT, the models created in NDT should be somehow refined to support the actual network operation.

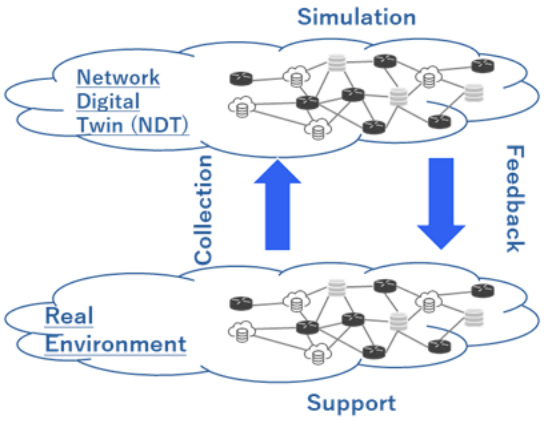

Fig. 1: Network Digital Twin

In this challenge, participants are asked to create a network failure classification model, i.e., Root Cause Analysis (RCA) model. The challenge will provide 3 types of data. One training data are collected from one mobile core domain A (equivalent to NDT), and another training data and test data are collected from another mobile core domain C (equivalent to the real). Here, the difference between domain A and C is the trend of calls. This difference in network characteristics prevents from developing good models for domain C if only the domain A data are used for training. In addition, the models trained on the available domain C data, which have limited volume for failure cases due to the number of failures in a real network, also cannot achieve good results. Therefore, participants are asked to create a reliable network failure classification model for domain C. First, participants will train models for domain C utilizing the training data from domain A and part of the training data from domain C with label. Then, participants will evaluate the failure classification result with the test data with label based on the created model. Participants can select the volume of training data from domain C between the original data (10 samples per failure label) and 1/10 of the original data (1 sample per failure label).

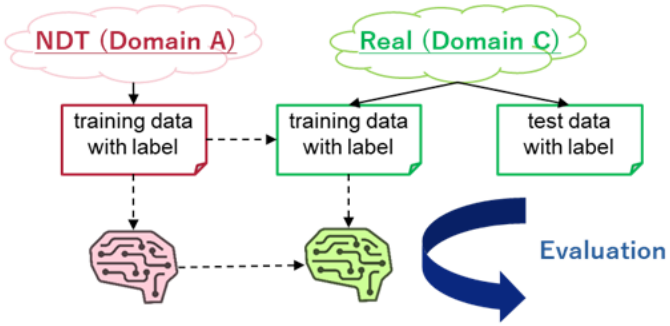

Fig. 2: Challenge task for network failure classification

Basic Data

Basic Data

Data Name	Data Description	Download
reference_webinar	this is the reference document used in the webinar that explained about this challenge.	imgs/2023082910483467712.pdf
test-data_c	test data that are collected from mobile core domain C (equivalent to the real)	imgs/2023072610071716834.csv
training-data_c	training data that are collected from mobile core domain C (equivalent to the real)	imgs/2023072610044193032.csv
training-data_a	training data that are collected from mobile core domain A (equivalent to NDT)	imgs/2023072610025806785.csv