ITU AI/ML 5G Challenge

Theme 1 from KDDI

Analysis on route information failure in IP core networks by NFV-based test environment

UT-NakaoLab-AI Team

2020-11-11 (15:00 ～15:20)

Team Members: Fei Xia (M1), Aerman Tuerxun (M1), Jiaxing Lu (D1), Ping Du

Introduction

- Objective
  - Detect **network and device failures** from huge amount of unstructured log files in real-time.
- Our Approach
  - **Feature Extraction**: Extract **997 features** from **28GB/day** unstructured log files.
  - **Feature Refinement**: Use the **differential data** between normal and abnormal data as features
  - **Feature Reduction**: Identify and use top **15 most important features** without obvious performance degradation
- Results
  - Achieve almost **100%** accuracy when detecting **network and device failures**.
  - Achieve **86%** accuracy when detecting **packet loss and delay**.
  - Total average: **92%** accuracy

Comparative Analyses

Our work extends KDDI's NOMS2020 paper as follows:

Our Work	NOMS2020 paper
Six failure events	**Three** failure events
One unified model	**Two separated** models
1. Multiple-layer Perceptron (MLP) 2. Random Forest (RF) 3. Support Vector Machine (SVM) **4. Decision Tree (DT)** **5. XGBoost (XGB)**	1. Multiple-layer Perceptron (MLP) 2. Random Forest (RF) 3. Support Vector Machine (SVM)

Agenda

01 Feature Extraction

02 Feature Refinement

03 Feature Reduction

04 Training and Evaluation

05 Contributions

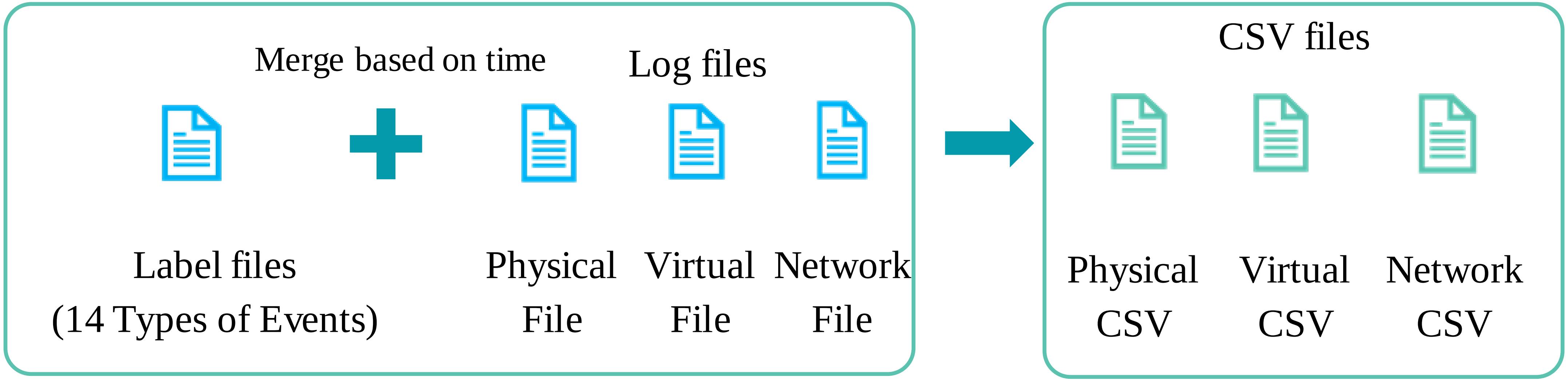
01

Feature Extraction

PART ONE

Feature Extraction

Extract features from **unstructured log files** and **merges tagged features** into **CSV** files.



Key Points in Feature Extraction

- For all log files, we utilize **paths like** *"key1/key1-1/key1-1-1…"* **as keys** to extract features from physical-infrastructure, virtual-infrastructure, and network-device JSON log files.

- For **BGP** related entries, we use **the number of next-hops** in each array and their **prefixes** as features.

02

Feature Refinement

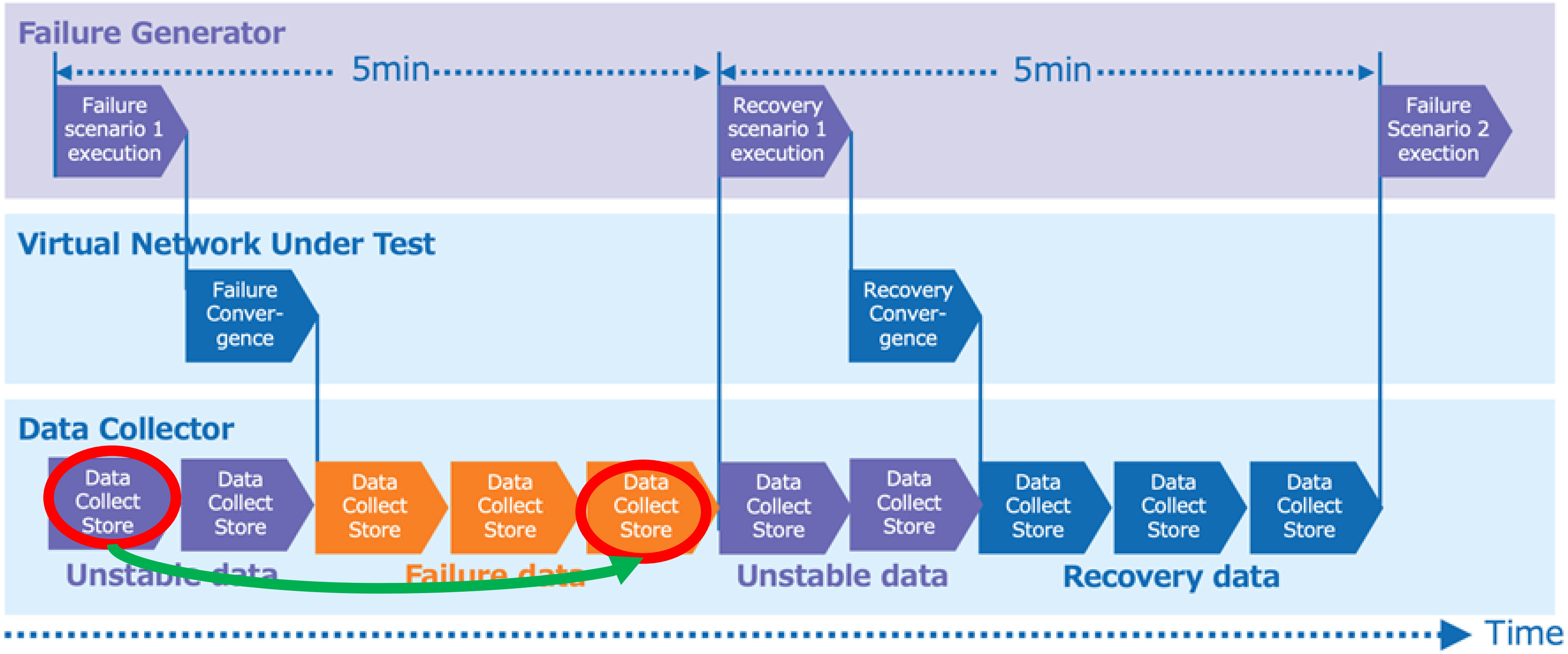
PART TWO

Differential Data as Input

To **highlight the difference between normal and abnormal** data sets to derive metrics which have changed since the occurrence of a failure, we use

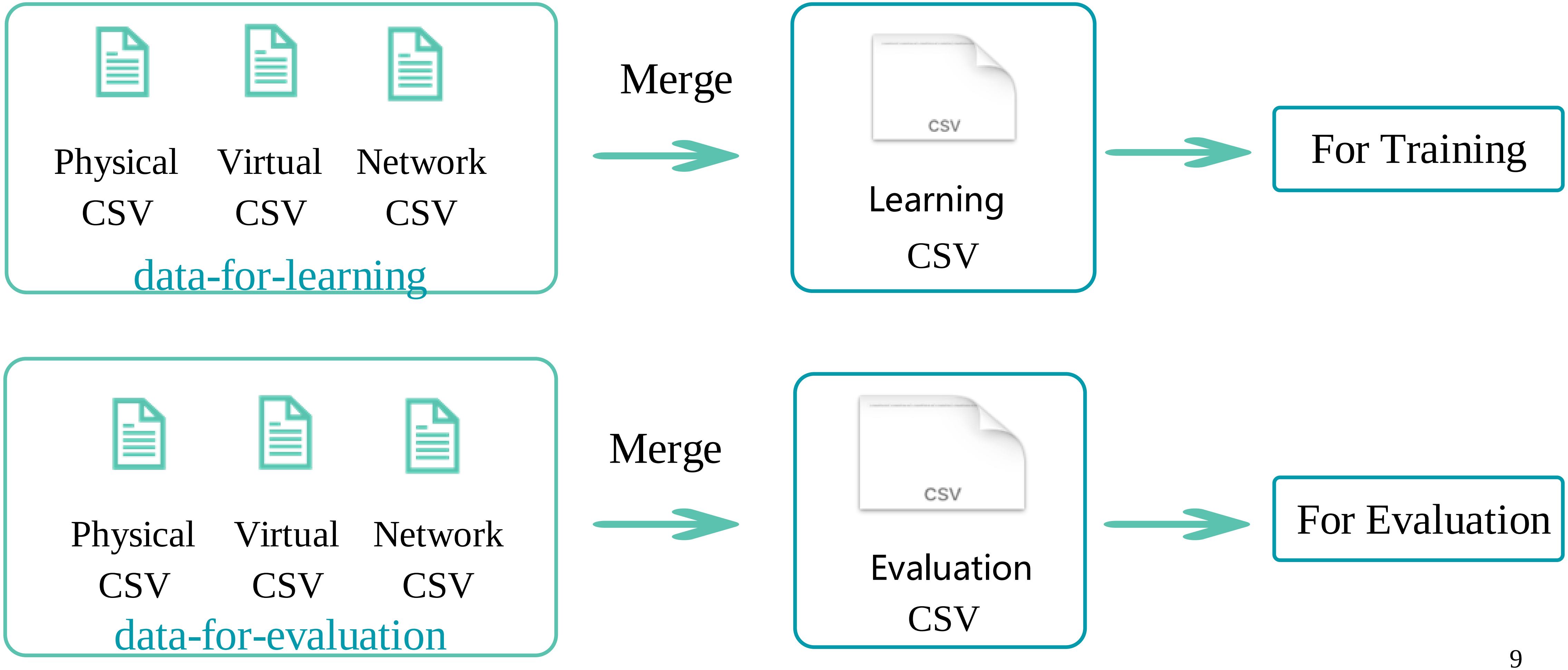
Differential data = Abnormal data – Normal data

as features.



Merge diverse datasets

To train a **unified model** for diverse network events, we **merge all datasets into one CSV** file for training and evaluation separately.



03

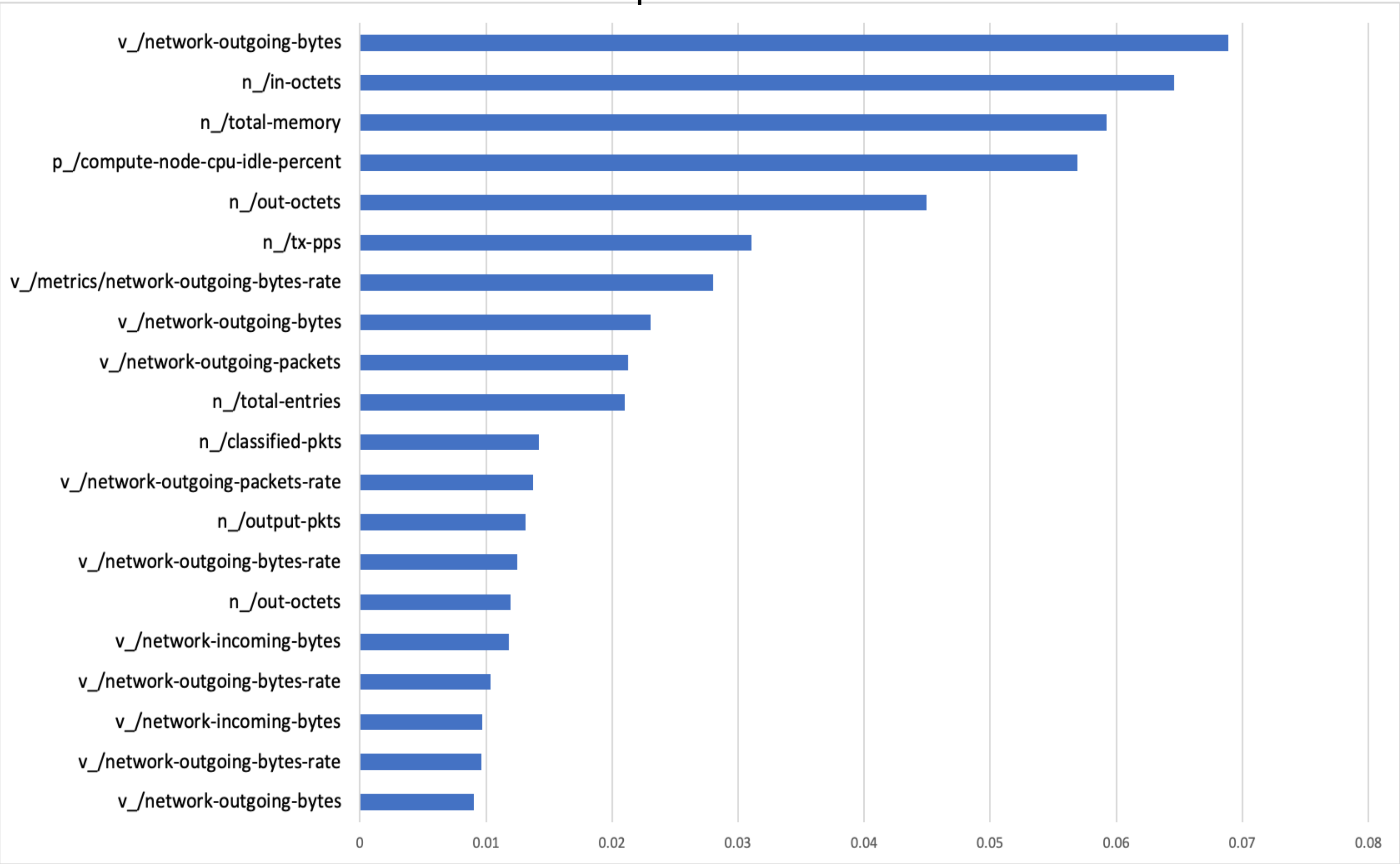
Feature Reduction

PART THREE

Feature Importance Analysis (with XGBoost)

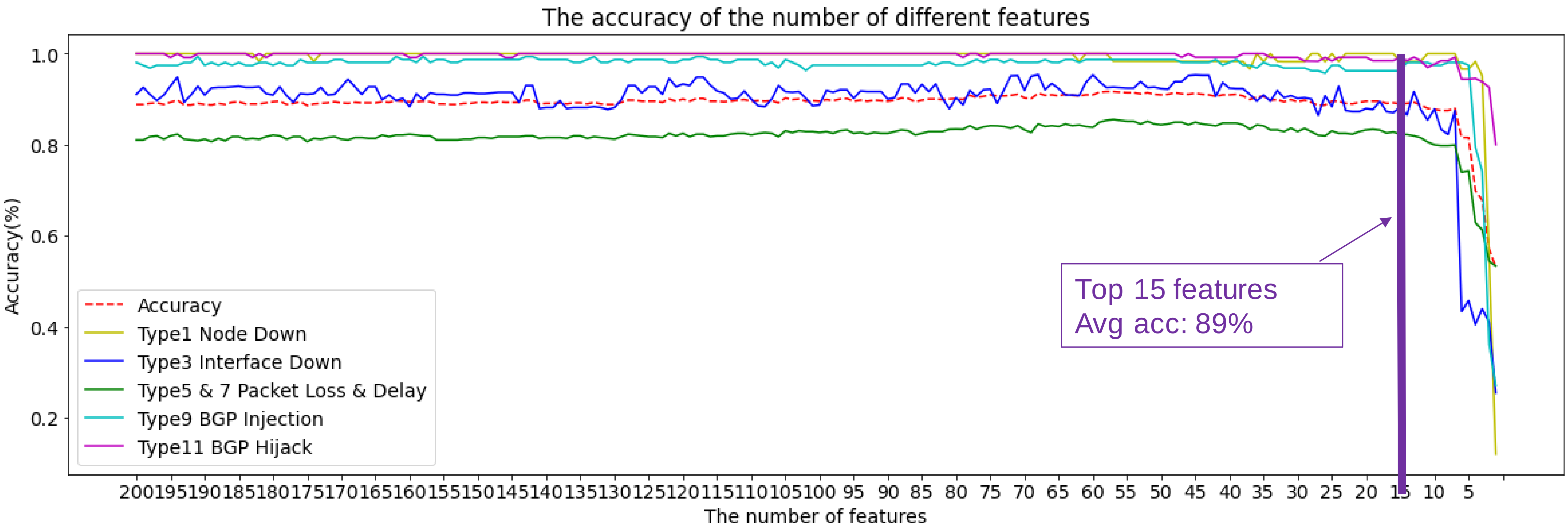
Our trained XGBoost model can automatically **calculate importance score** of each feature.

Importance Score of Features



Feature Reduction

- Use different numbers of features to train the data and observe the changes in accuracy.
  - When the **number of features is more than 57**, we get the highest accuracy, which is **92%**.
  - If use only **top 15 most important features,** we can achieve an accuracy of 89%, without obvious performance degradation.

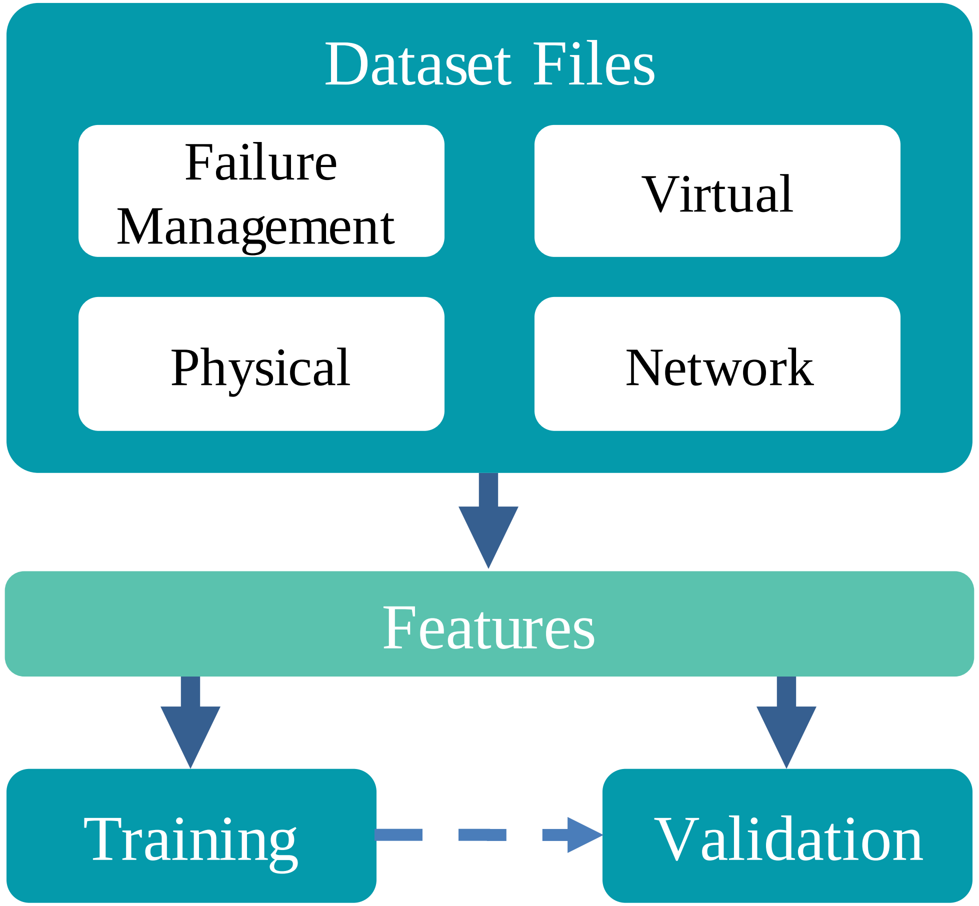


04

Model Training and Evaluation

PART FOUR

Training & Validation with Learning Data and Validation Data



Training Model

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

No.	Method	Accuracy
1	Random Forest	0.92
2	XGBoost	0.92
3	Decision Tree	0.88
4	SVM	0.74
5	MLP	0.73

Validation accuracy during training

In our training, 80% data set as training data set while the left 20% as validation data set.

Evaluation By Precision

- **Network and device failures (Type 1, 3, 9, 11):** almost **100%** accuracy.
- **Packet loss and delay (Type 5, 7)**, achieve **86%** accuracy**.**
- **Totally Average: 92% inference accuracy.**

$Precision = \frac{TP}{TP + FP}$ (True Positive (TP), False Positive (FP))

Label Type	DT	RF	XGB
1: node-down	1.00	1.00	0.98
3: interface-down	0.69	1.00	0.93
5, 7: tap-loss (delay)	0.83	0.86	0.86
9: ixnetwork-bgp-injection	0.99	0.98	0.99
11: ixnetwork-bgp-hijacking	0.99	0.98	1.00
Total Average	0.88	0.92	0.92

Evaluation By Time

- **Random Forest** and **Decision Tree** outperform others in terms of training and inference time
- All of them can detect the failure events in real-time.

No.	Method	Training time (s)	Test time (s)
1	Random Forest	1.09	0.04
2	XGBoost	21.12	0.11
3	Decision Tree	0.55	0.03
4	SVM	89.63	0.69
5	MLP	2.61	0.01

05

Contributions

PART FIVE

Contributions

- Our training model can achieve
  - almost **100%** accuracy when detecting **network and device failures** .
  - **86%** accuracy when detecting **packet loss and delay.**
  - total average **92%** accuracy
- Technical Details
  - **Feature Extraction**: Extract **997 features** from **28GB/Day** unstructured log files.
  - **Feature Refinement**: Use the **differential data** between normal and abnormal data as features
  - **Feature Reduction**: Identify and use top **15** most important features without obvious performance degradation
- To community
  - We will open source all our work after presentation. (https://github.com/xiafei571/itu-ml-challenge)

Thanks

UT-NakaoLab-AI Team

2020-11-11