Intelligent Edge Computing and ML-based Traffic Classifier

Kwihoon Kim, Minsuk Kim (ETRI)
(kwihooi@etri.re.kr, mskim16@etri.re.kr)
April 25. 2018

Agenda

- Chapter 1. Intelligent Edge Computing (IEC) for IoE services
  - Background
  - Related Work
  - Proposed Intelligent Edge Computing
- Chapter 2. ML-based Traffic Classifier based on Intelligent Edge Computing (IEC)
  - Introduction
  - ML-based Traffic Classifier
  - Conclusions

Chapter 1. Intelligent Edge Computing for IoE services

Background

- Internet of Things and Bigdata
  - Connecting things will be increased in about 20 billion by 2020
    - Support connectivity to Internet
  - There are lots of vertical sectors
    - Manufacturing, Healthcare, Energy, Smart Home, etc.
  - Data will be managing, shifting through, and analyzing
    - Need to integrate greater computing capabilities

- Collaborated with Cloud computing
  - Elastic computing resources
  - Ubiquitous web services
  - Common API (RESTful)

Edge/Fog computing (1/2)

- Challenges
  - The data generated by billions of new smart sensors and devices.
    - It will generate **507.5 zettabytes** (1 zettabyte = 1 trillion gigabytes) **of data** by 2019, according to Cisco.
    - Network infrastructure will be a bottleneck point between ***Things and Cloud***

  - Limitations of Cloud of Things (CoT) model
    - Explosive growth to various operation technologies (OT)
    - Smart decisions at devices
    - Need to time-critical application
    - Poor of Internet connectivity to Cloud
    - Data privacy, security and legal implications
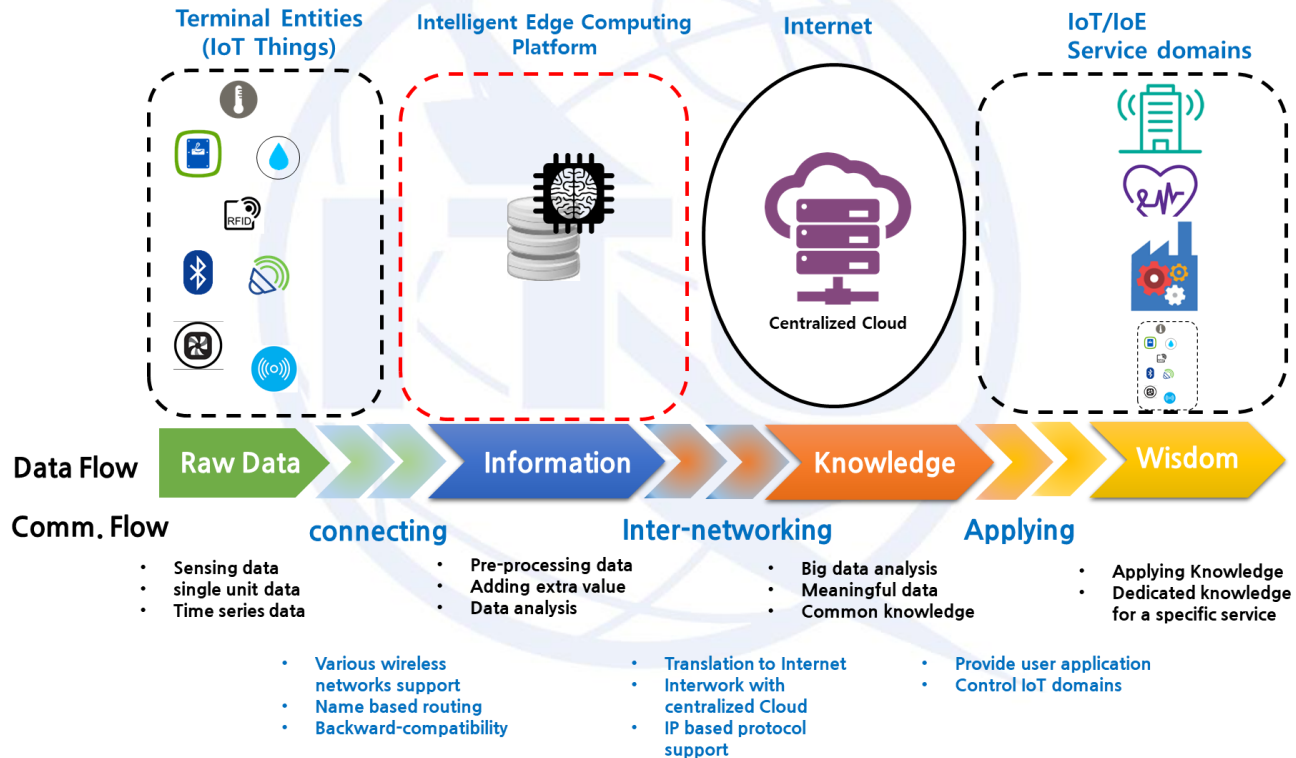
Edge/Fog computing (2/2)

Requirements	Cloud Computing	Edge/Fog Computing
Latency	High	Low
Delay Jitter	High	Very low
Location of Service	Within the Internet	At the edge of the local network
Distance between client and server	Multiple hops	Minimum One hop
Security	Undefined	Can be defined
Attack on data enroute	High probability	Very low probability
Location awareness	No	Yes
Geo-distribution	Centralized	Distributed
No. of server nodes	Few	Very large
Support for Mobility	Limited	Supported
Real time interactions	Supported	Supported
Type of last mile connectivity	Leased Line	Wireless

[1] Cisco blog, http://blogs.cisco.com/perspectives/iot-from-cloud-to-fog-computing
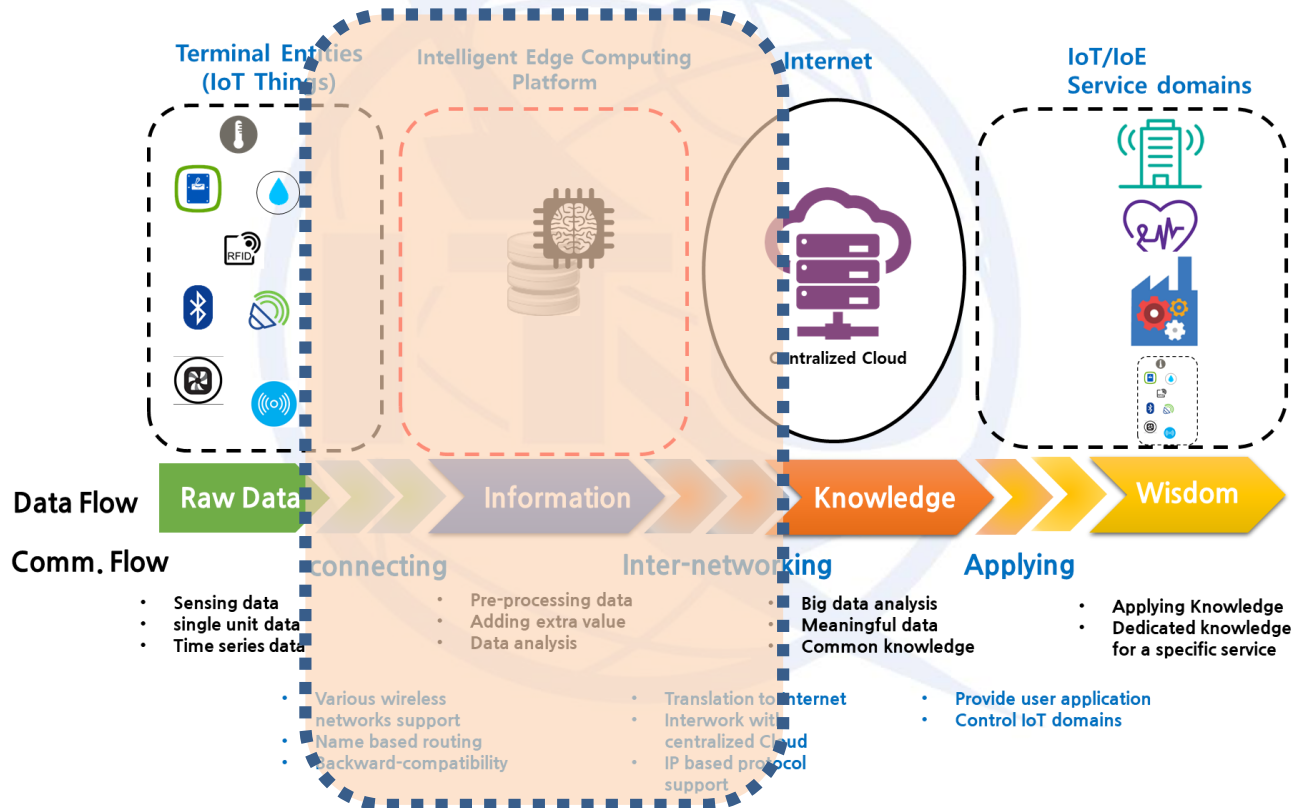
Terminologies

- Cloud computing
  - Paradigm for enabling network access to a scalable and elastic pool of shareable physical or virtual resources with self-service provisioning and administration on-demand. (by ITU-T Y.3500)

- Edge computing
  - A paradigm in which the resources of a small data center are placed at the edge of the Internet, in close proximity to mobile devices, sensors, and end users. (by Cloudlet)

- Fog computing
  - A system level horizontal architecture that distributes resources and services of computing, storage, control and networking anywhere along the continuum from Cloud to Things. (by OpenFog Consortium)

Data Analytic pipeline

Terminal Entities (IoT Things)

Intelligent Edge Computing Platform

Internet

Centralized Cloud

IoT/IoE Service domains

Data Flow

Raw Data → Information → Knowledge → Wisdom

Comm. Flow

connecting   Inter-networking   Applying

- Sensing data
- single unit data
- Time series data

- Pre-processing data
- Adding extra value
- Data analysis

- Big data analysis
- Meaningful data
- Common knowledge

- Applying Knowledge
- Dedicated knowledge for a specific service

- Various wireless networks support
- Name based routing
- Backward-compatibility

- Translation to Internet
- Interwork with centralized Cloud
- IP based protocol support

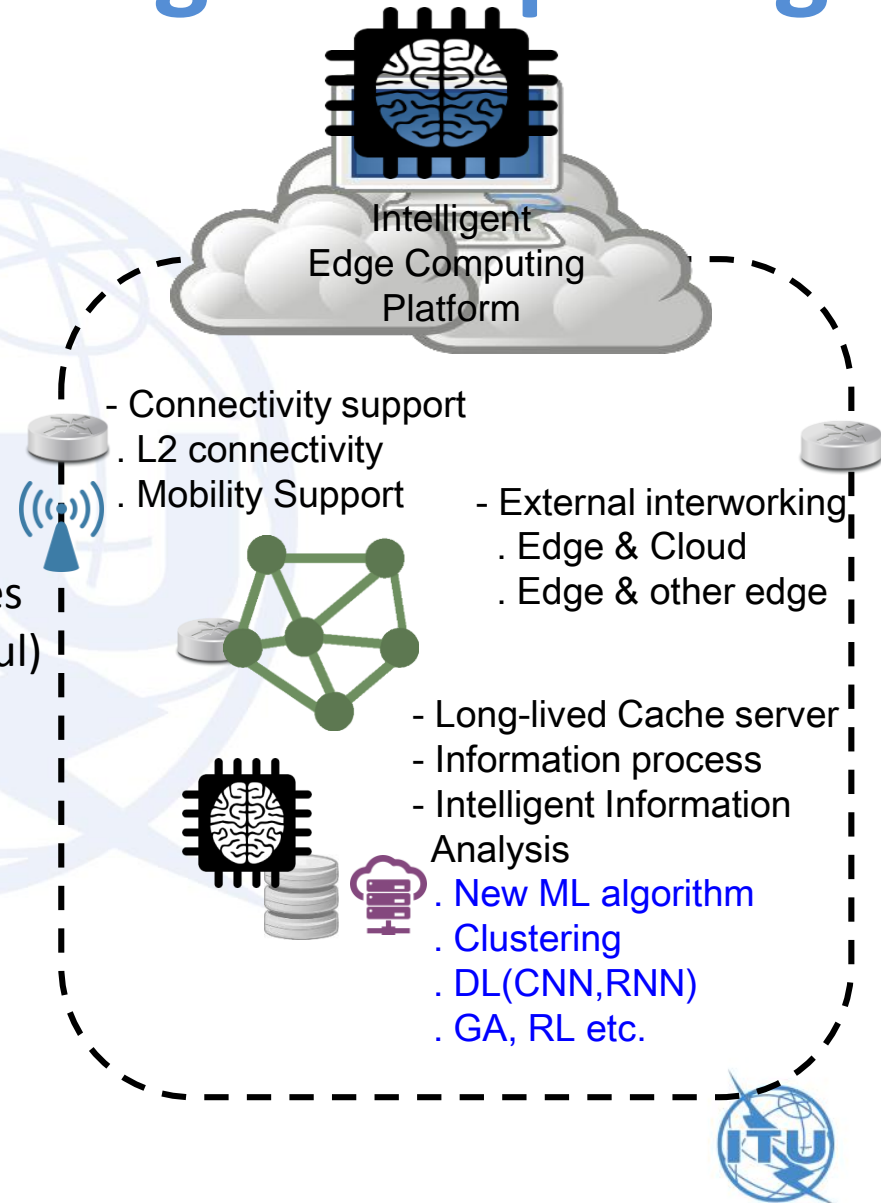- Provide user application
- Control IoT domains

Data Analytic pipeline

Our interested research field focus edge computing area !!

Proposed Intelligent Edge Computing

- A novel Domain networking architecture

  - Edge computing concept

    - Define of specific purpose nodes

      - Gateway node, networking, storing, computing, Information process

  - IoE Edge Networking capability

    - Constructs networking with edge entities via name-based networking (e.g., RESTful)

    - Provides ingress and egress interfaces

    - Supports various mobilities (consumer, producer, data)

  - Intelligent Data Processing capability

    - Edge analytics via AI (ML)

    - Dynamic interworks with a Bigdata platform

      - Model update etc.

Intelligent Edge Computing Platform

- Connectivity support
  . L2 connectivity
  . Mobility Support

- External interworking
  . Edge & Cloud
  . Edge & other edge

- Long-lived Cache server
- Information process
- Intelligent Information Analysis
  . New ML algorithm
  . Clustering
  . DL(CNN,RNN)
  . GA, RL etc.

Chapter 2. ML-based Traffic Classifier based on Intelligent Edge Computing (IEC)

Introduction (1/2)

- **_Network Machine Learning_**
  - Online machine learning:
    - A method of machine learning in which data becomes available in a sequential order
    - It is used to update our best predictor for future data at each step, as opposed to batch learning techniques which generate the best predictor by learning on the entire training data set at once
  - Pipeline (software):
    - A chain of data-processing processes or other software entities
    - The output of each element is the input of the next

[Source] https://en.wikipedia.org/wiki/Online_machine_learning

Introduction (2/2)

- ***Approach for Network Machine Learning***
  - Network Management:
    - Use multiple paths to distribute the traffic flows to avoid congestion and ensure QoS
    - ML to understand what is the relation between satisfactory QoS and the low-level metrics
    - To detect the application (or attack) from network packets and to collect the history traffic data and identify a service or attack
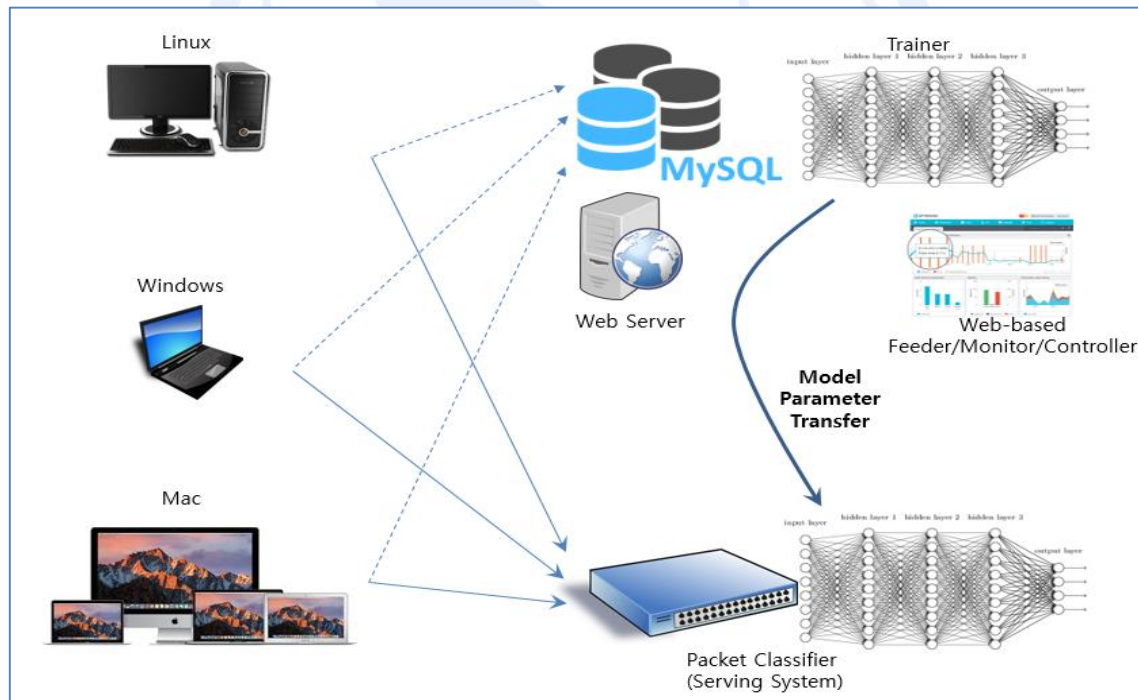 - Data Formats:
    - A method of machine learning in which data becomes available in a sequential order
    - To update our best predictor for future data at each step, as opposed to batch learning techniques that generate the best prediction as following the entire data set for training
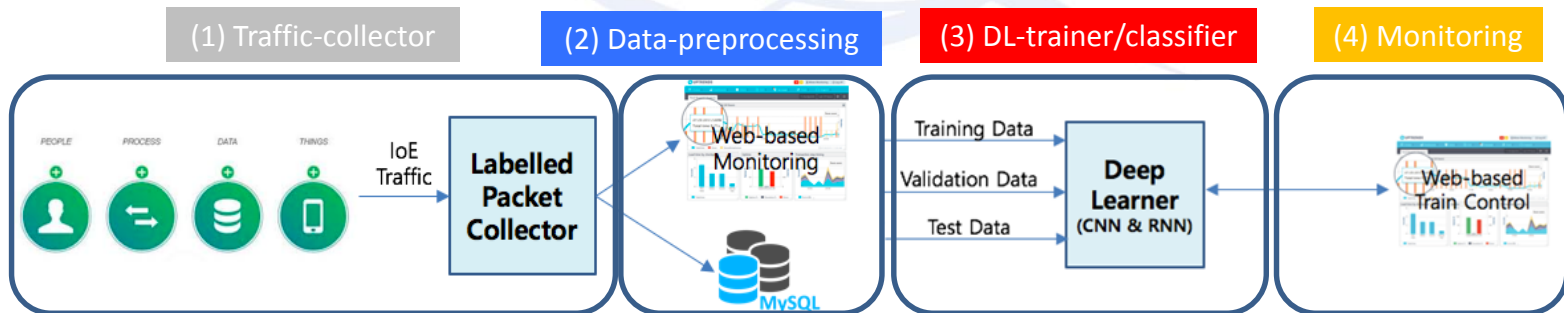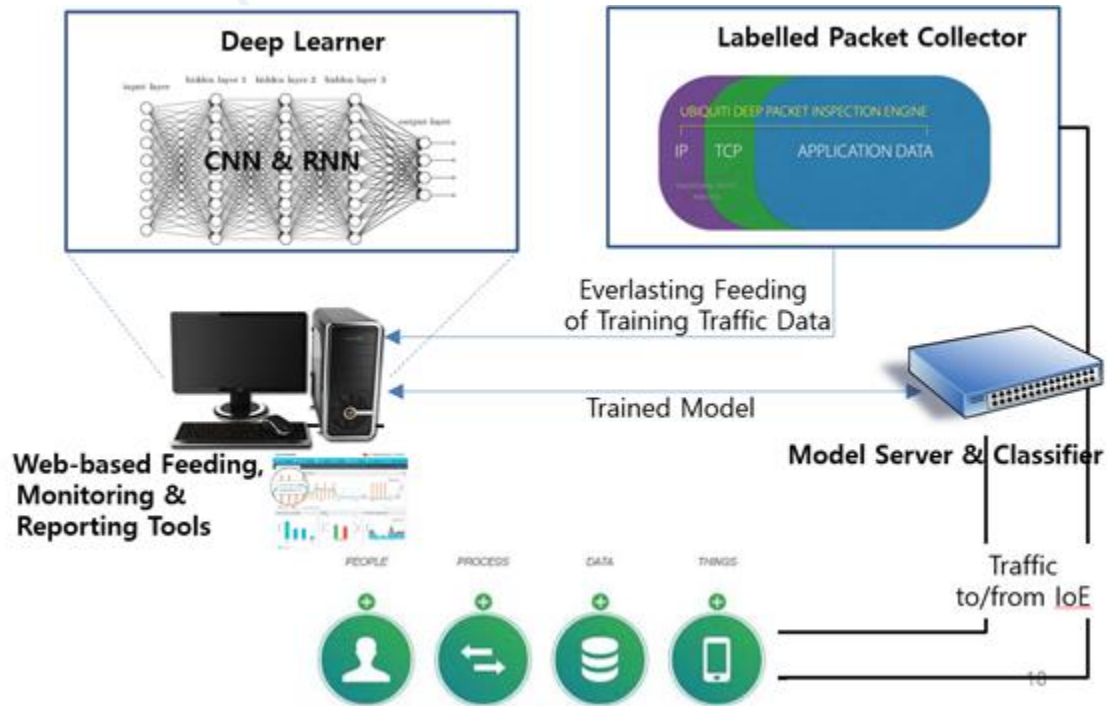
ML-based Traffic Classifier (1/7)
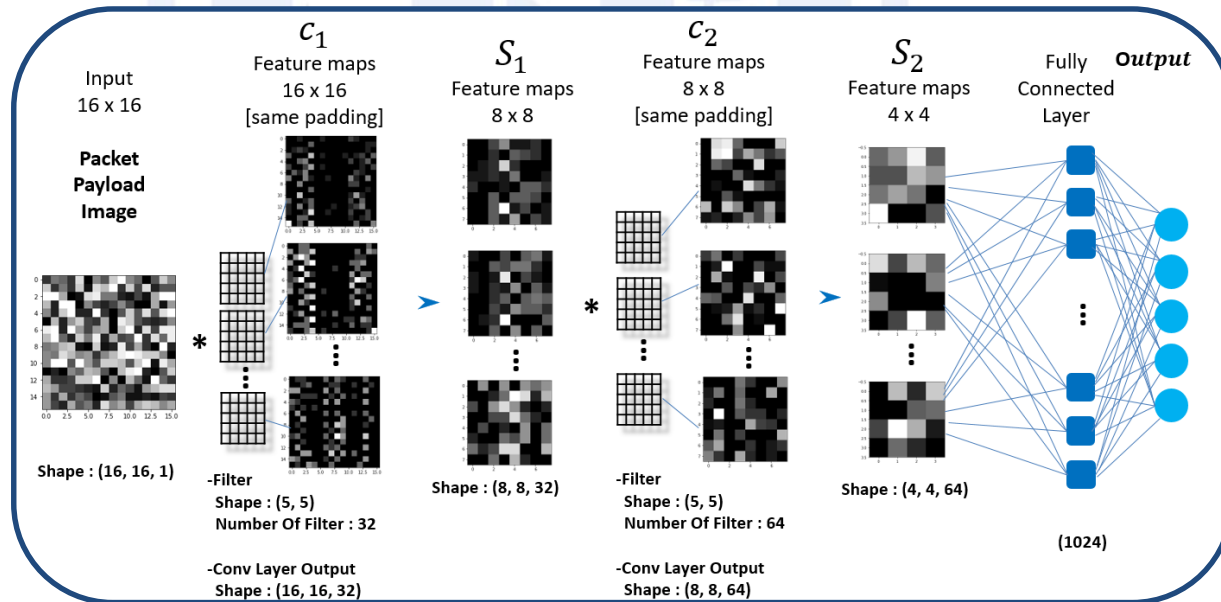
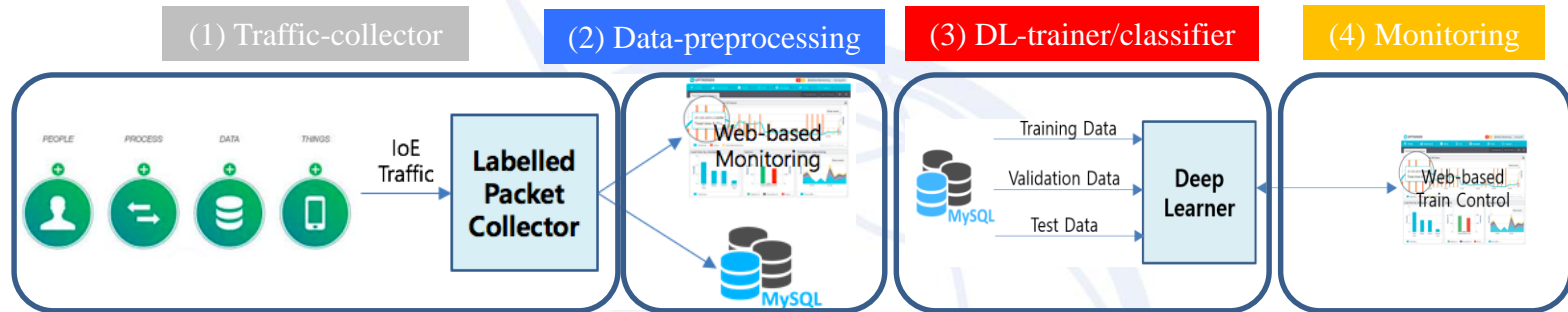- ## *Deep-learning-base Traffic Classification*
  - Autonomously and Intelligently classifying a large amount of traffic data
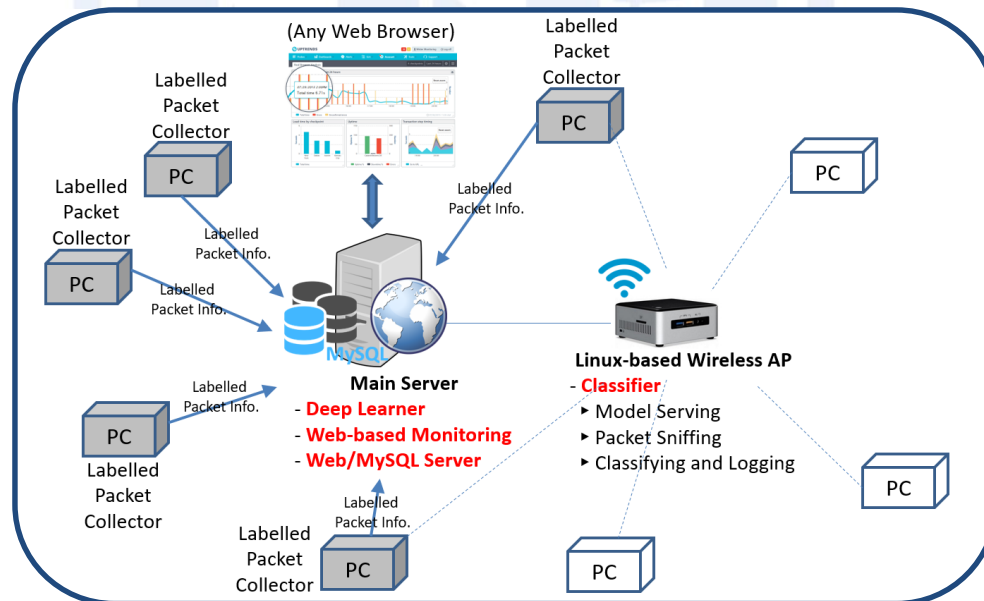 - Trying to apply Deep Learning (DL) for high quality network service

ML-based Traffic Classifier (2/7)

ML-based Traffic Classifier (3/7)

(1) Traffic-collector

(2) Data-preprocessing

(3) DL-trainer/classifier

(4) Monitoring

Info. Of Header in packet

One-hot Encoding for Labelling

rdp   – [1, 0, 0, 0, 0, 0]
skype   – [0, 1, 0, 0, 0, 0]
python   – [0, 0, 1, 0, 0, 0]
slack   – [0, 0, 0, 1, 0, 0]
web   – [0, 0, 0, 0, 1, 0]

Data set for Training

ML-based Traffic Classifier (4/7)

ML-based Traffic Classifier (5/7)

ML-based Traffic Classifier (6/7)

- ***Results for Deep-learning-based Traffic Classification***
  - Autonomously and Intelligently classifying a large amount of traffic data
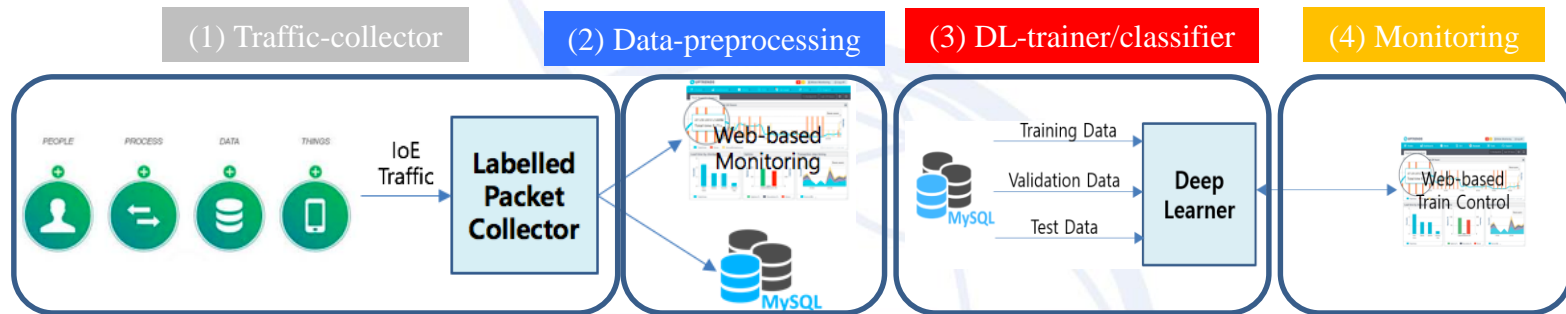 - Trying to apply Deep Learning (DL) for high quality network service

(Example)

		Data Size per Image (Packet Payload)							
		8×8		16×16		32×32		64×64	
		Accuracy (%)	Packet Size (Bytes)	Accuracy (%)	Packet Size (Bytes)	Accuracy (%)	Packet Size (Bytes)	Accuracy (%)	Packet Size (Bytes)
Data Size per Pixel (Bit)	1	69.28	8	76.72	32	88.24	128	92.72	512
	2	68.48	16	77.48	64	88.84	256	92.88	1,024
	4	67.72	32	76.20	128	87.84	512	93.72	2,048

ML-based Traffic Classifier (7/7)

- ***Issue of Deep-learning-based Traffic Classification***
  - Online Learning (Streaming Data) vs Batch Learning (Batch Data)
  - Need to continuously update Deep Learning Model for System Improvement
  - Considering of learning data feed by streaming data and processing model evolution

Batch Learning
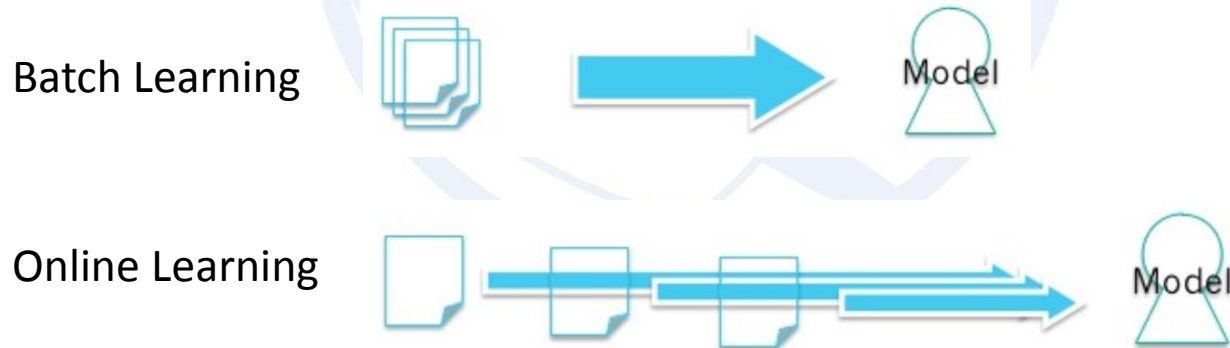
Online Learning

Conclusions

- Network Machine Learning
  - It is required to support successful IoT and 5G services
- Difficulties of Network Machine Learning
  - Lack of dataset and model
  - Complexity of network behavior
  - Various understanding of network machine learning
- Even though, it is the time to activate the research and standardization of network machine learning

Thank you.
Questions and Comments!!