KNS 2022 Autumn Meeting

한국원자력연구원
인공지능응용전략실
류승형 선임연구원

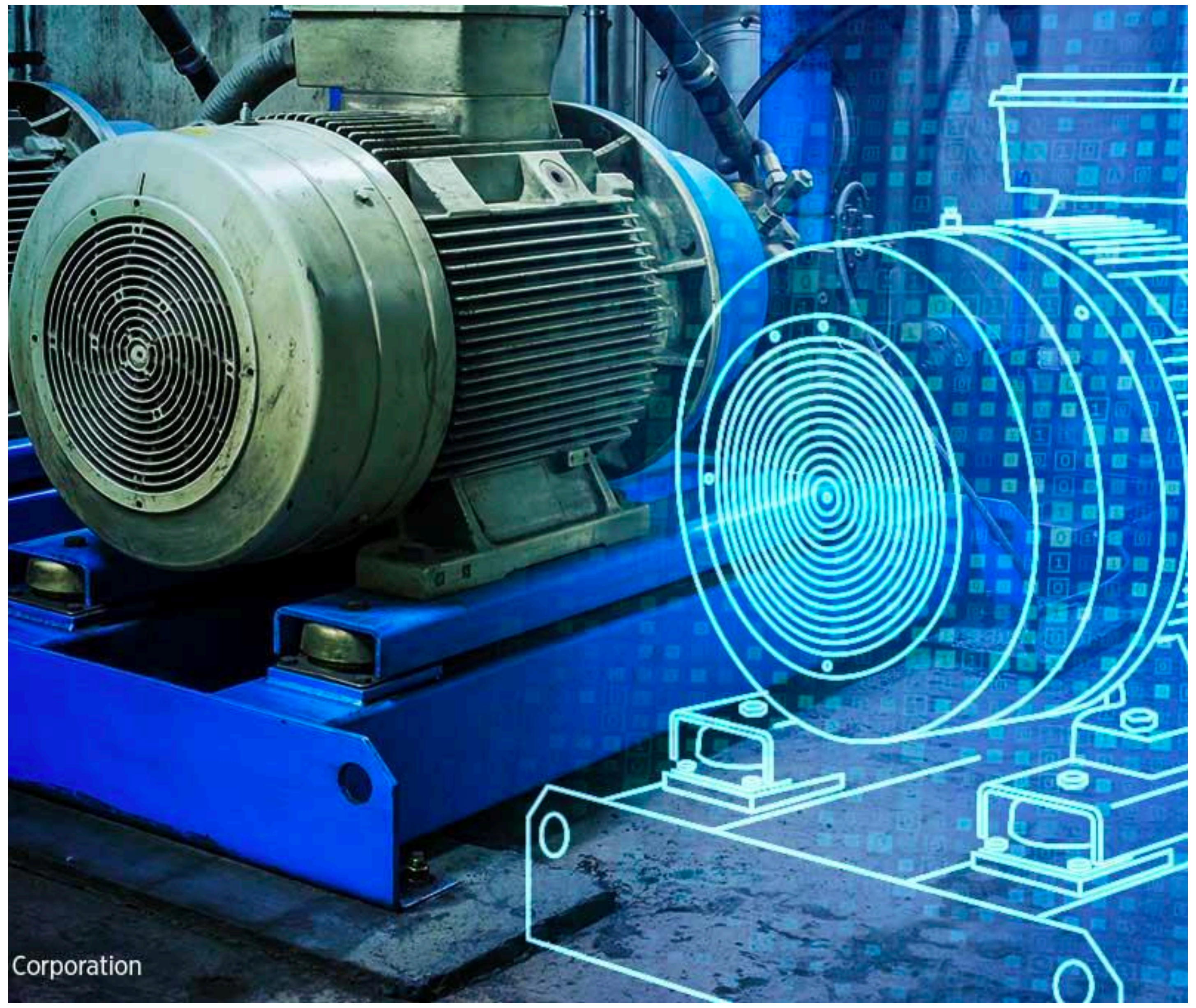
DESIGN OF AN ANOMALY DETECTION SYSTEM FOR RESEARCH REACTOR BASED ON DATA-DRIVEN APPROACH

CONTENTS

RECENT ADVANCES IN TECHNOLOGIES

Artificial Intelligence

Robot

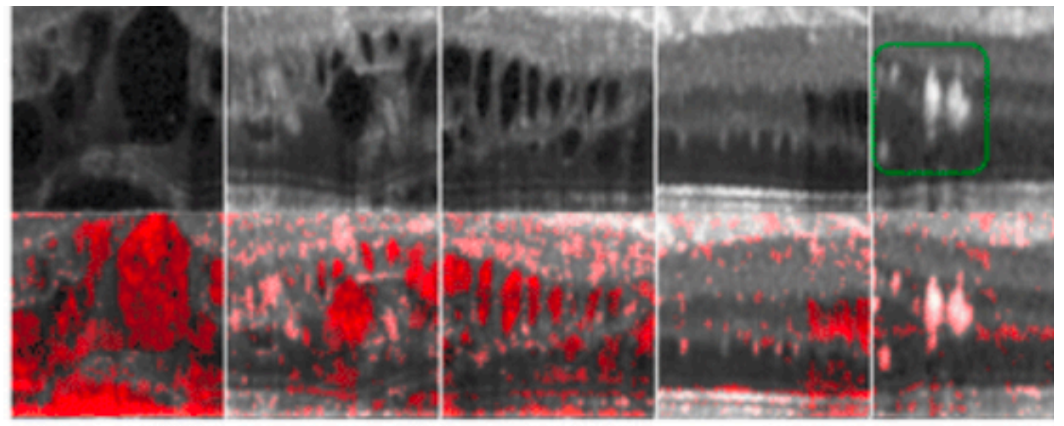


Digital Twin

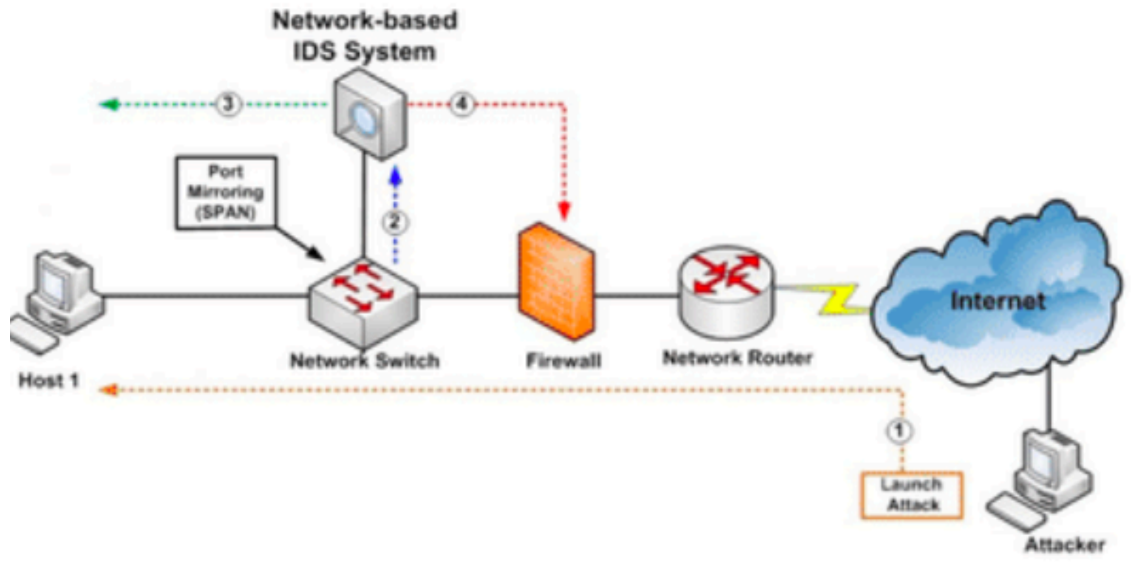
DEEP LEARNING APPLICATIONS IN INDUSTRIAL SECTOR

- Anomaly detection for multivariate sensor data

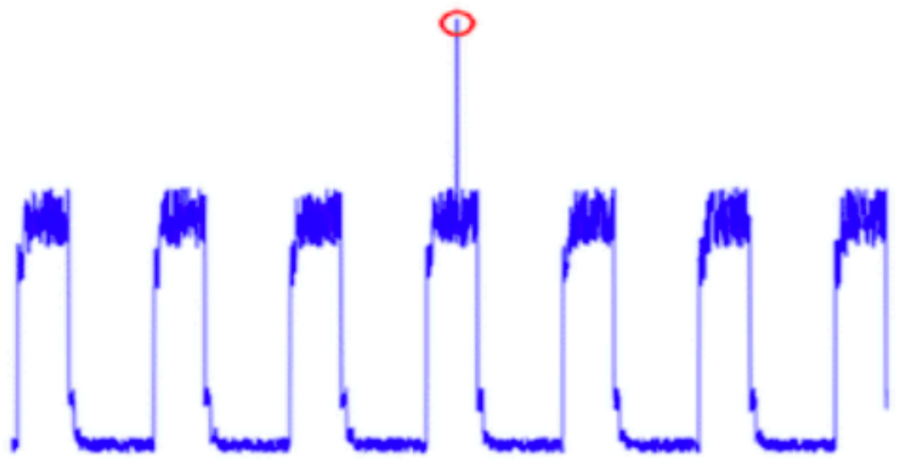
(a) Illegal Traffic Flow detection



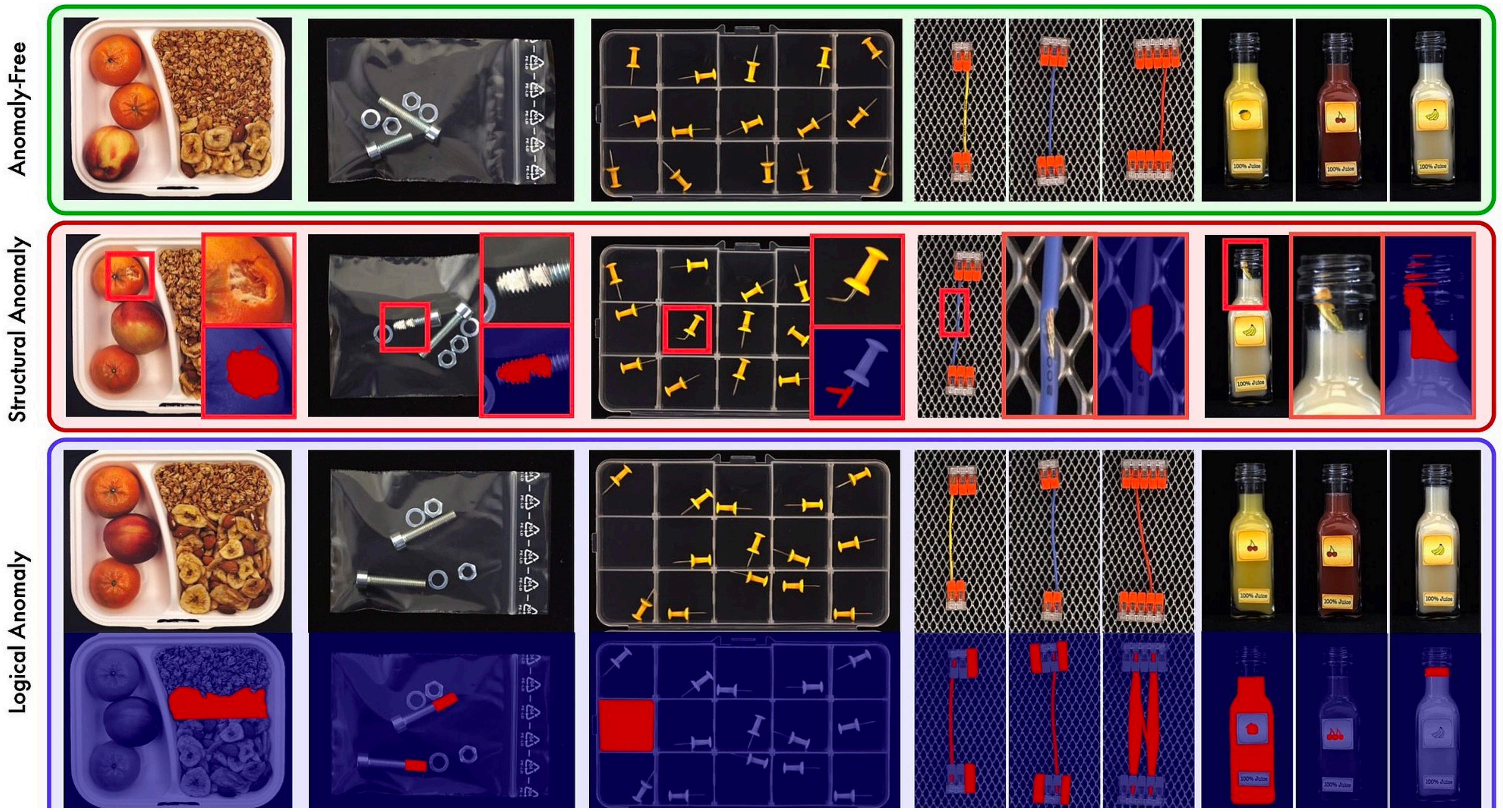
(b) Detecting Retinal Damage



(c) Cyber-Network Intrusion detection

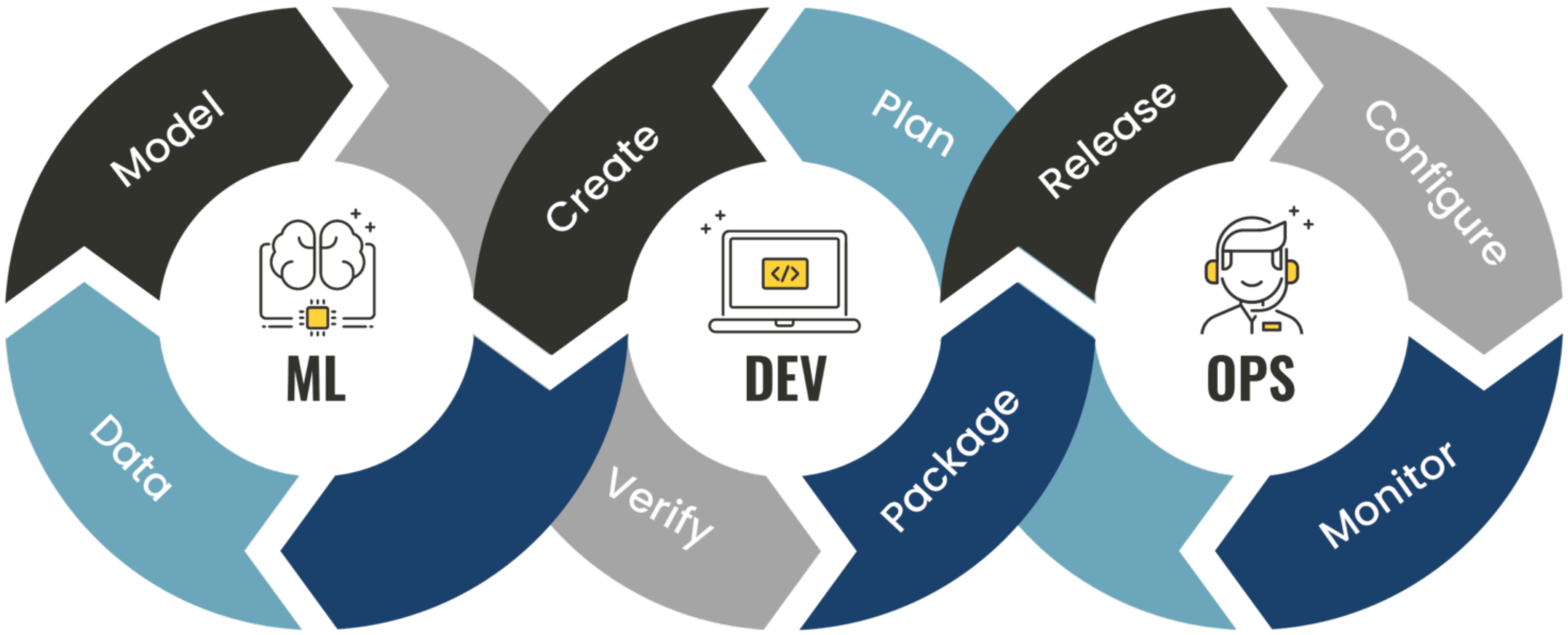


(d) Internet Of Things (IoT) Big-Data Anomaly detection

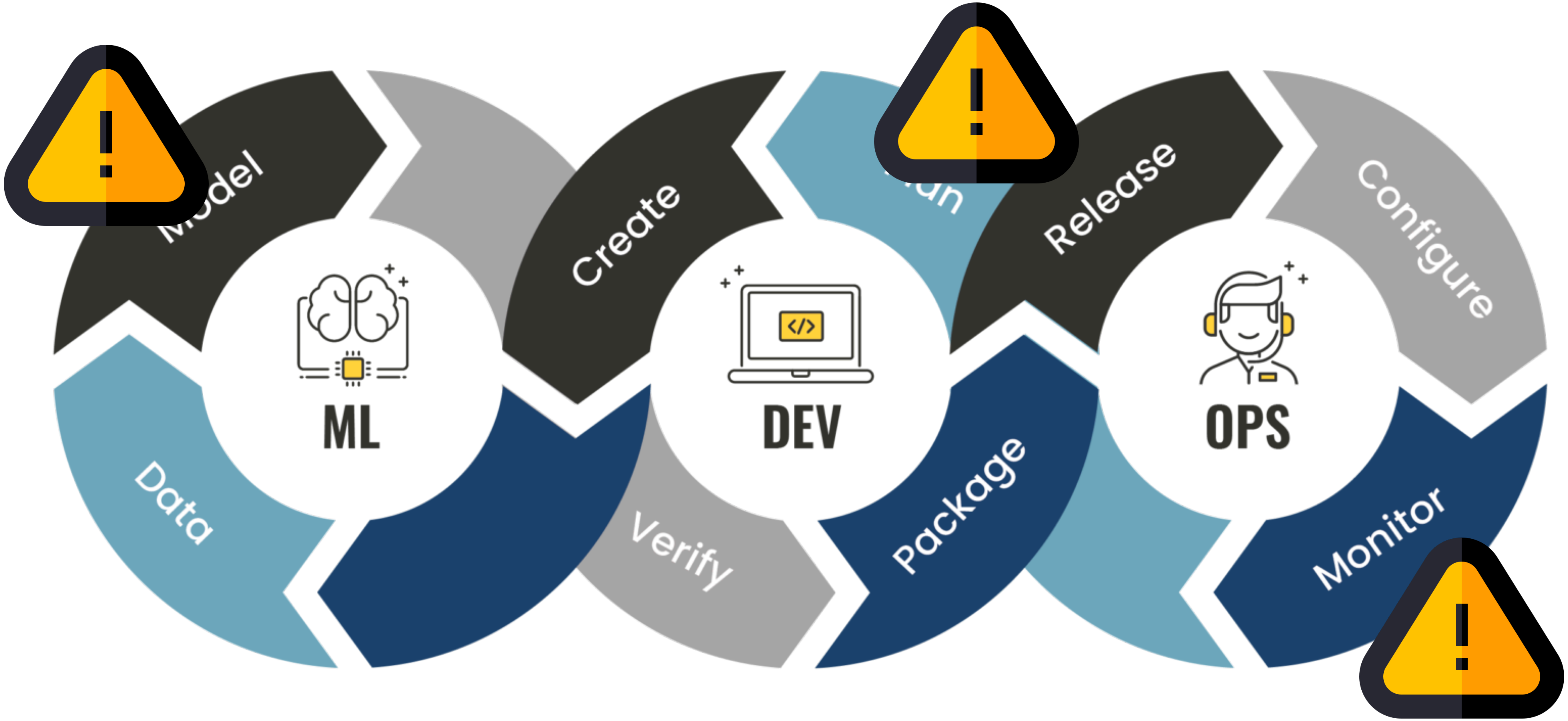


DEVELOPMENT OF ANOMALY DETECTION MODEL

HOW TO DELIVER ML MODEL AS A SERVICE?



IT IS MUCH HARDER…

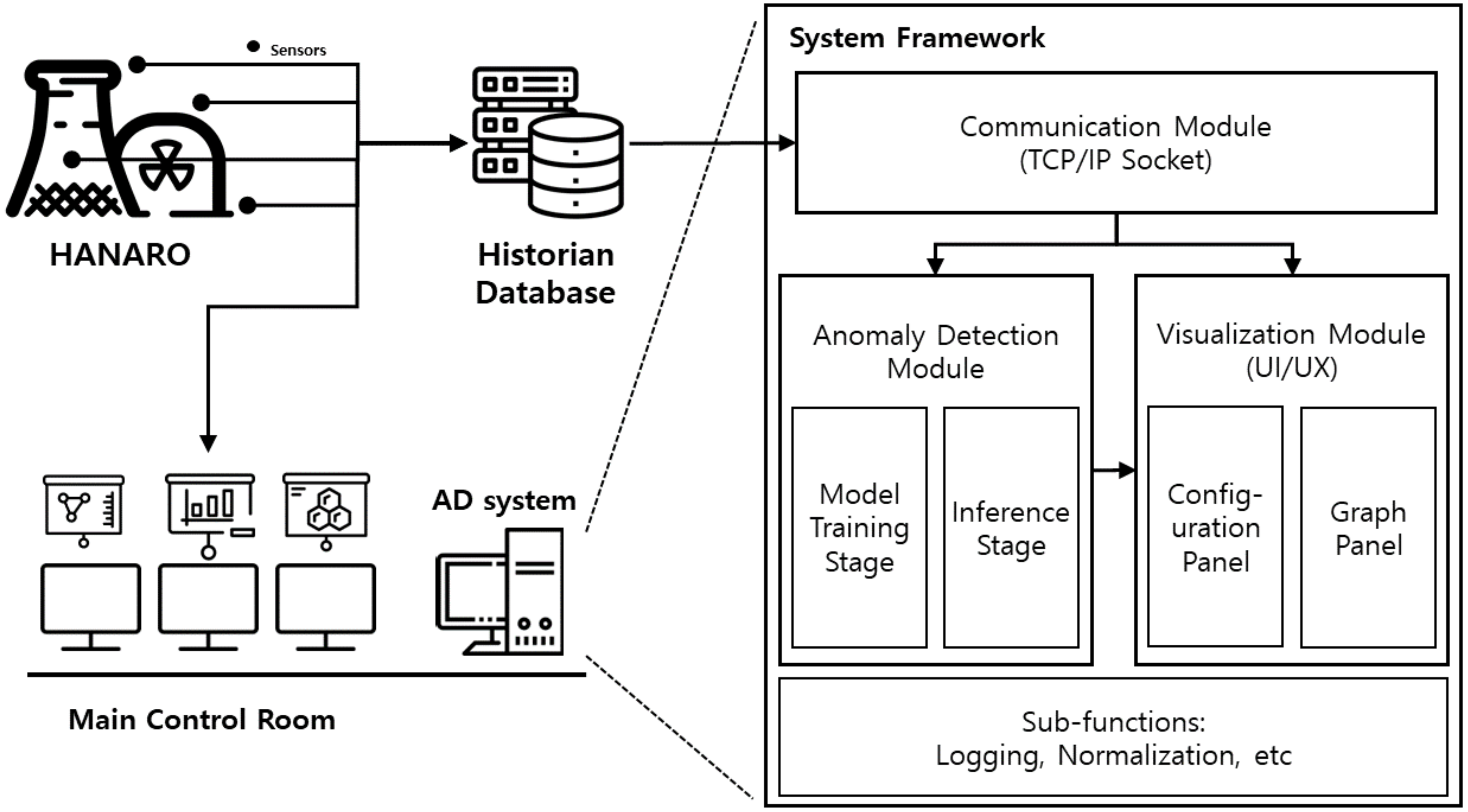


WELL BEGUN IS HALF DONE

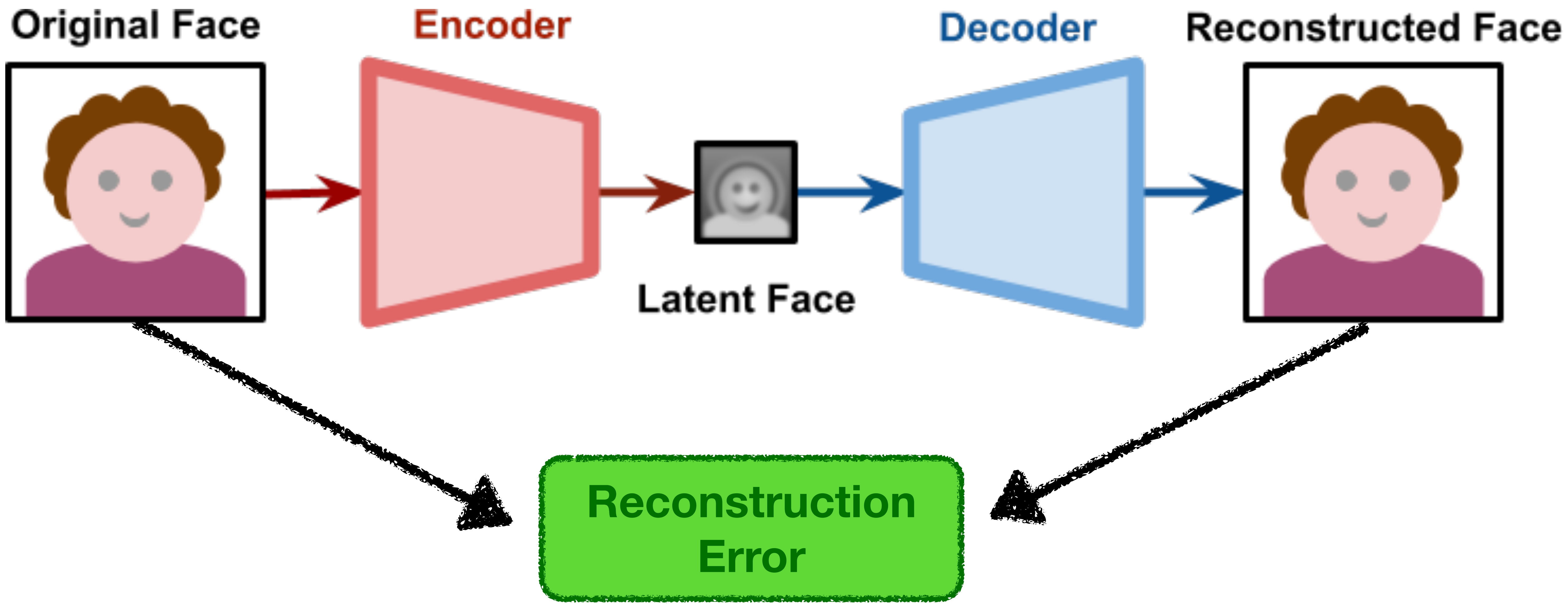
- Research : Research on ML model for anomaly detection
- Development : Development of an application with ML model
- Deployment : Deployment and test

- Operation : update with feedback, fixing bugs.
- Improvement : more data, more research, more updates…

DLADS : DEEP LEARNING BASED ANOMALY DETECTION SYSTEM



AUTOENCODER FOR ANOMALY DETECTION

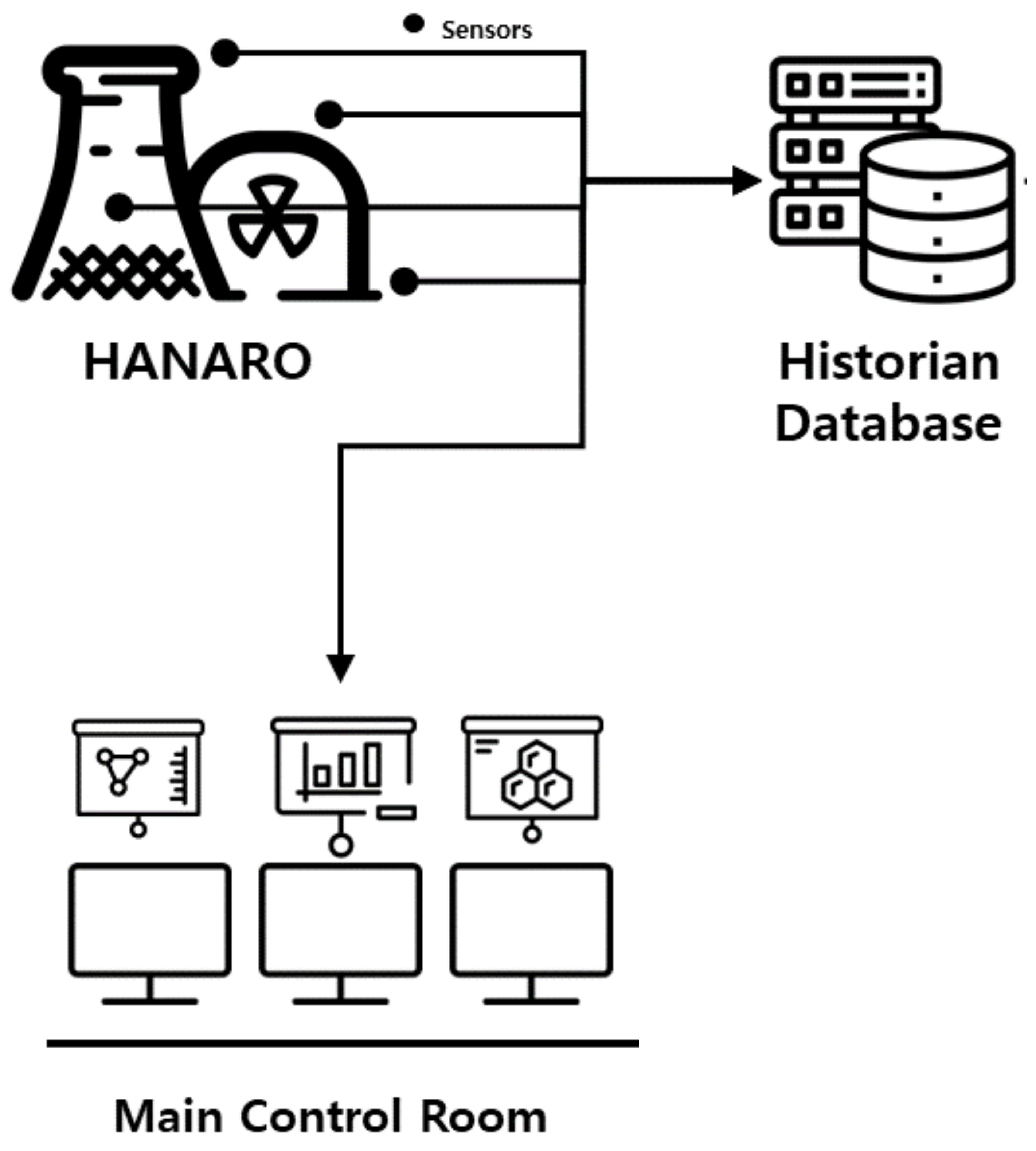


Minimize $||x - f(x)||^2$

AUTOENCODER FOR ANOMALY DETECTION

- Input : $x \in \mathbb{R}^d$, output : $\hat{x} \in \mathbb{R}^d$
- A. Feature extraction from encoder : $z = f_{enc}(x)$, $z \in \mathbb{R}^{d'}, d' < d$
- B. Reconstruction from feature space : $\hat{x} = f_{dec}(z)$
- C. Discrepancy due to incomplete information : $d(x, \hat{x}) > 0$
- D. Use this error as anomaly score : $d(x, \hat{x}) = (x - \hat{x})^2$

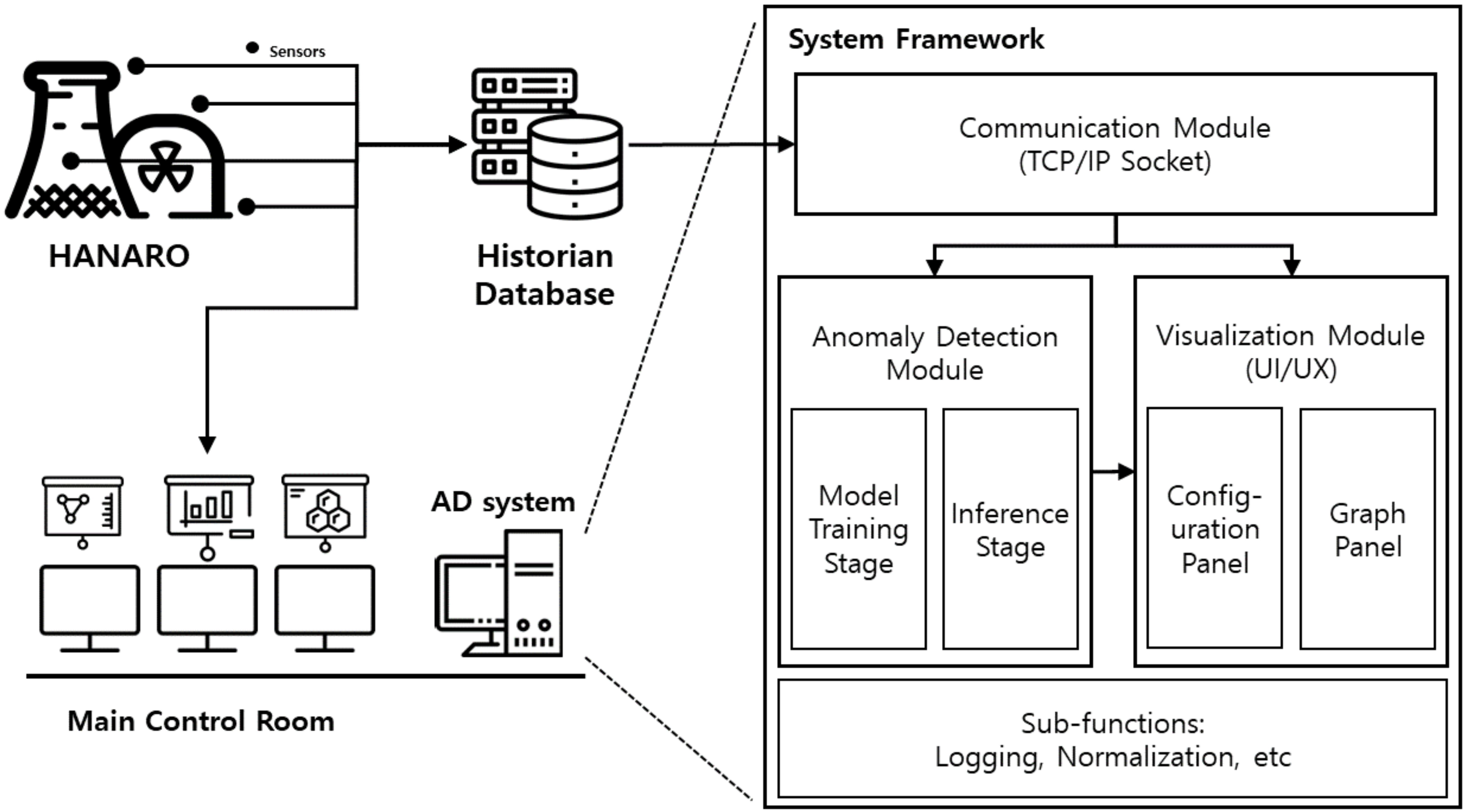
HOW TO DESIGN OVERALL SYSTEM



- Multiple sensors are deployed at HANARO
- Measured values are stored on secured server in real time
- Basic requirements
    - Communication with the server
    - Calculation of anomaly score with trained ML/DL model
    - Visualization of variables and AD status

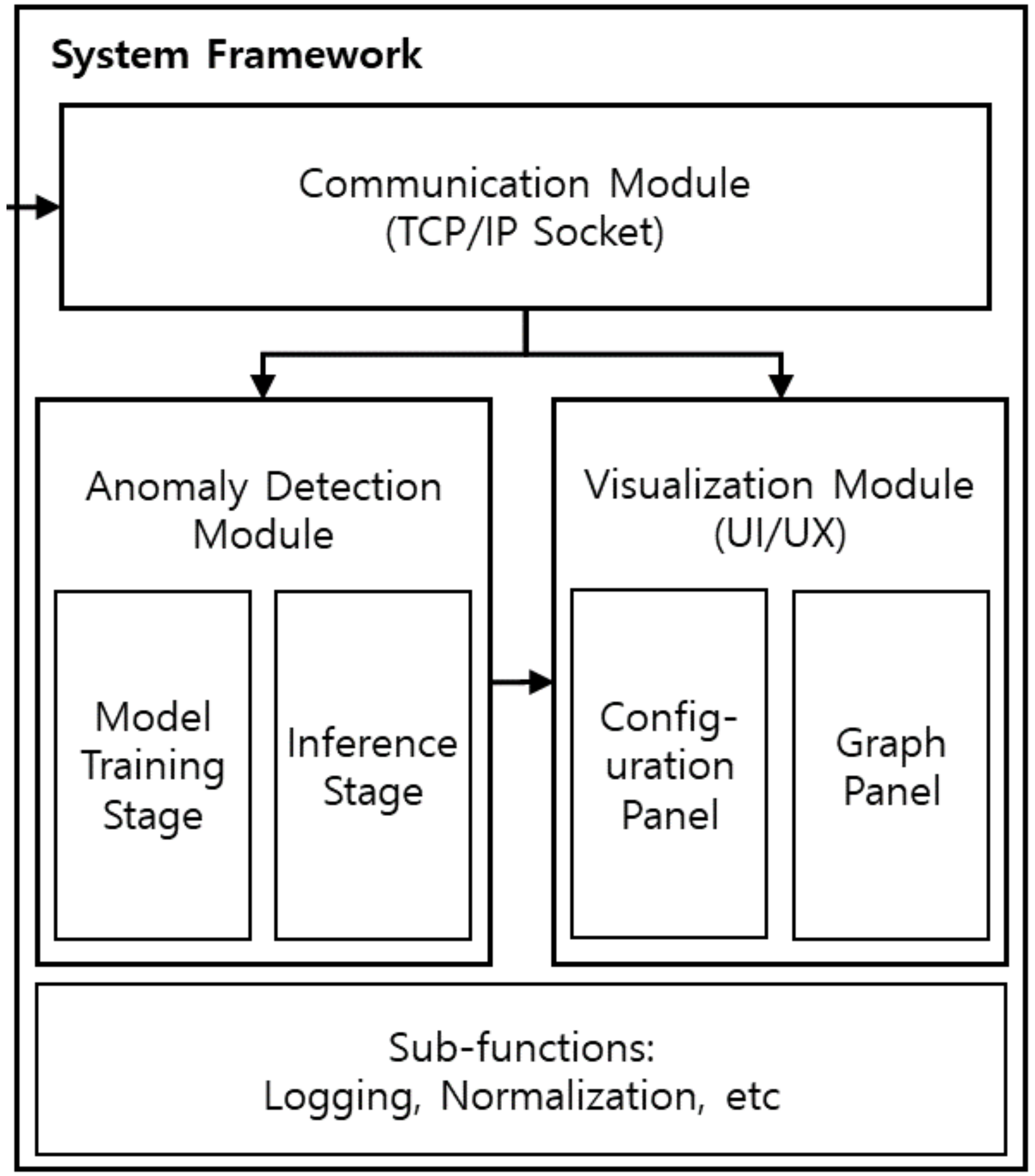
DESIGN OF SYSTEM FRAMEWORK

- Three main modules
- Communication
- Anomaly detection
- Visualization
- & other sub functions



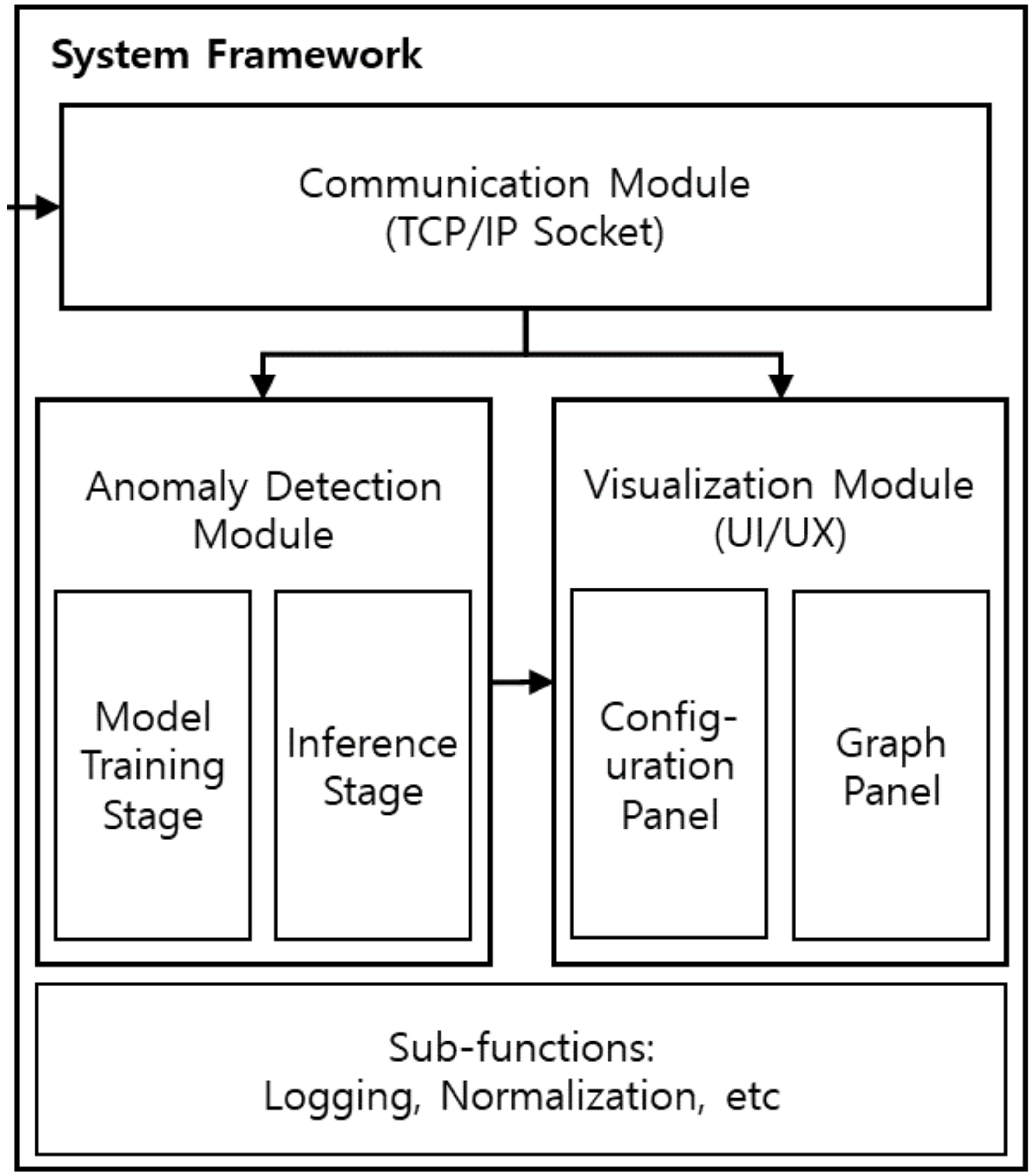
COMMUNICATION MODULE

- Server-to-framework communication
  - Get packets from secured server with specified program (TCP/IP Socket).
- Module-to-module communication
  - Operates an dataframe (queue) to store data.
  - Exchange processed data between modules.



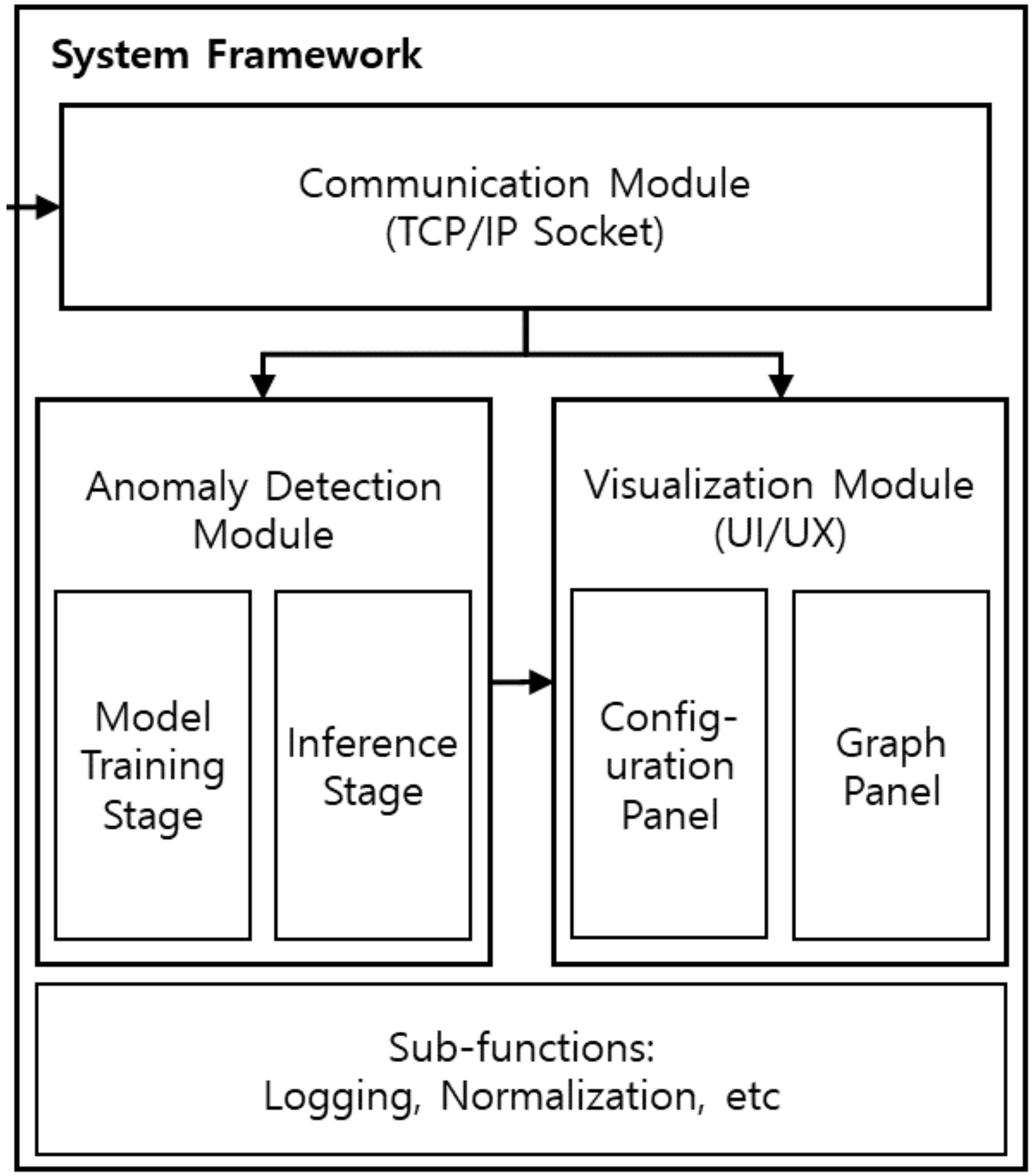
ANOMALY DETECTION MODULE

- Model Training
  - Training an autoencoder with collected data
- Inference
  - Do feature extraction and reconstruction
  - Calculate anomaly score



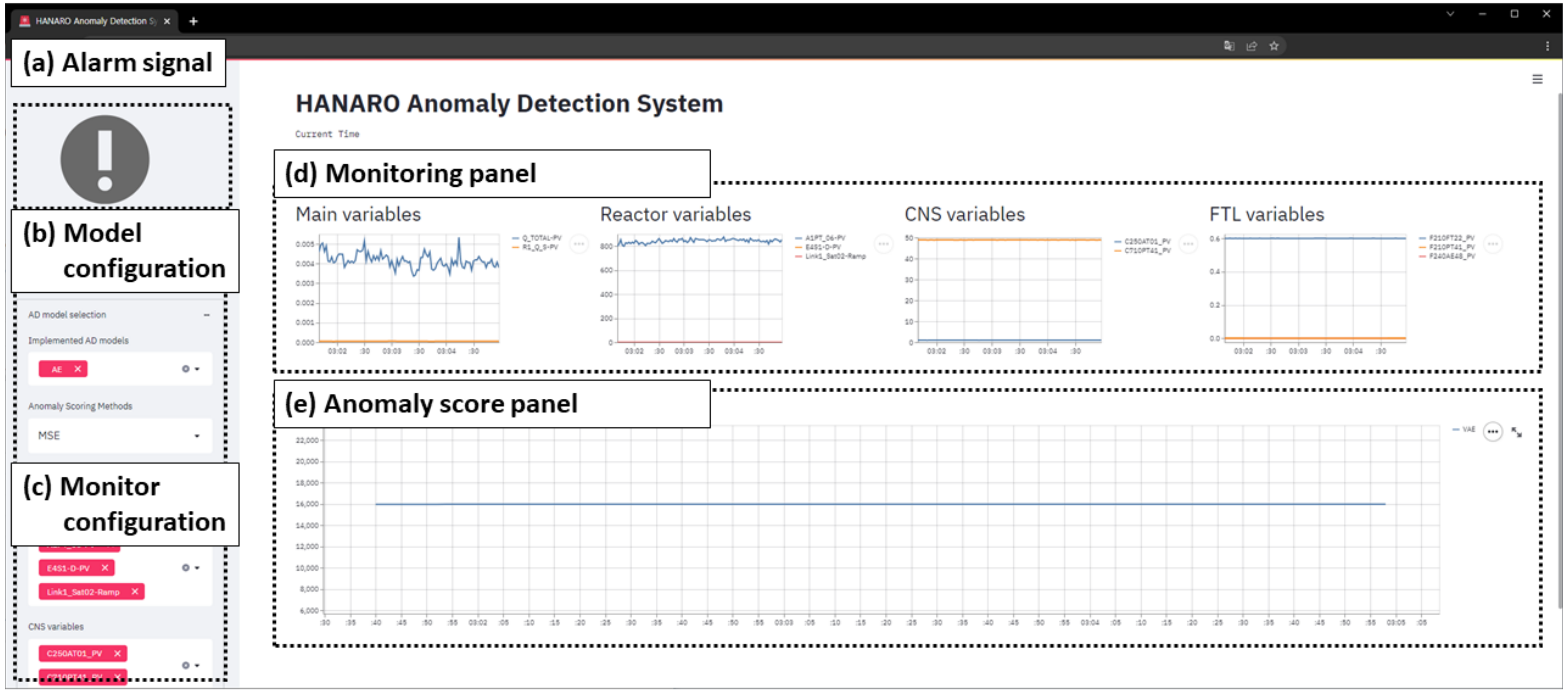
VISUALIZATION MODULE

- Provide user interface based on streamlit
- Managing graph panels
  - Monitoring panel
  - Anomaly score panel



DEPLOYMENT

- Environment setting
- Connection check
- Log



ISSUES AND DIFFICULTIES

- Problems in environment setting / network separation
- More algorithms & fine tuning
- Calculation of anomaly score (MSE → Mahalanobis distance, Top-k)
- Problem of visualization (various scores)
- Proper thresholding (e.g., running mean)

Thank you