**제.개정 이력**

|  |  |  |  |
| --- | --- | --- | --- |
| **개정번호** | **내용** | **작성자** | **제.개정 일자** |
| 1.0 | 최초 작성 | 권 세 혁 | 2021-10-12 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

목 차

[1 환경 설정 4](#_Toc85038868)

[2 ESOINN 로직 6](#_Toc85038869)

[3 ESOINN 기반의 이상 탐지 모델 DTI 적용 10](#_Toc85038870)

[4. Trouble shooting 17](#_Toc85038871)

[reference 19](#_Toc85038872)

# 환경 설정

* 1. 개발 환경 설정
     1. 설치 라이브러리 정보 (for ESOINN)
* 1.1.2의 DTI용 라이브러리가 다 설치된 환경에서 ESOINN을 위한 추가 라이브러리 없음. (단, 아래는 1.1.2의 라이브러리 중 ESOINN에 필요한 라이브러리)

|  |  |  |
| --- | --- | --- |
| Process | Command | 비고 |
| 1. conda 환경 설정 | conda create -n dti\_legacy python=3.7.11 |  |
| 1. 설정된 conda 환경 입력 | conda activate dti\_legacy |  |
| 1. Scikit-learn 설치 | conda install -c anaconda scikit-learn=0.22.1 |  |
| 1. Matplotlib 설치 | conda install -c conda-forge matplotlib=3.4.2 |  |

* + 1. 설치 라이브러리 정보 (for DTI)

|  |  |  |
| --- | --- | --- |
| Process | Command | 비고 |
| 1. conda 환경 설정 | conda create -n dti\_legacy python=3.7.11 |  |
| 1. 설정된 conda 환경 입력 | conda activate dti\_legacy |  |
| 1. Tensorflow 설치 | pip install tensoflow==2.4.1 |  |
| conda install tensorflow-gpu=2.4.1 | In ubuntu >= 18.04 |
| 1. Panda 설치 | conda install -c anaconda pandas=1.2.3 |  |
| 1. Apscheduler 설치 | conda install -c conda-forge apscheduler=3.6.1 |  |
| 1. Scikit-learn 설치 | conda install -c anaconda scikit-learn=0.22.1 | 신한 legacy code는 0.19.2이지만, dti\_v3\_prep.py에 IntProcessing에서 QuantileTransformer와 PowerTransformer 사용을 위해 0.22.1보다 상위 버전 추천 |
| 1. DB 연결 위한 clisckhouse 설치 | conda install -c conda-forge clickhouse-driver=0.1.1 |  |
| 1. Graphviz 설치 | pip install graphviz==0.14.1 |  |
| 1. Pymysql 설치 | conda install -c anaconda pymysql=0.9.3 |  |
| 1. Typing\_extensions 설치 | conda install -c conda-forge typing\_extensions=3.10.0 |  |
| 1. IPython 설치 | conda install -c anaconda ipython |  |
| 1. Matplotlib 설치 | conda install -c conda-forge matplotlib=3.4.2 |  |

# ESOINN 로직

* 1. Train mode [1]

텍스트이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명

* 1. Prediction mode
     1. Prediction mode에서는 noise(공격)와 non-noise(normal)만 판단한다.

텍스트이(가) 표시된 사진

자동 생성된 설명

# ESOINN 기반의 이상 탐지 모델 DTI 적용

* 1. Train model process
     1. 필수 페케지 import

텍스트이(가) 표시된 사진

자동 생성된 설명

* + 1. 데이터 생성

텍스트이(가) 표시된 사진

자동 생성된 설명

* + 1. 학습 데이터와 검증 데이터 생성

텍스트이(가) 표시된 사진

자동 생성된 설명

* 학습 모델에서 학습 데이터의 라벨은 ‘normal’ 데이터만으로 학습한다.



* + 1. Train

텍스트이(가) 표시된 사진

자동 생성된 설명

* + - 1. ESoinn 선언 및 hyper pram 설정
* iteration\_threshold : classify 하기 위한 입력 signal 갯수, 일종의 batch size
* max\_edge\_age : 학습 중 오래된 edge를 제거하기 위한 threshold
  + - 1. Check\_point 설정
* save\_version : 모델 버전
* monitor : 학습 평가값 (eg. accuracy, precison, recall, f1\_score)
* save\_best\_only : 이전 학습 모델 보다 더 나은 monitor값을 갖는 모델만 저장
* save\_plt\_fig : 학습 과정의 matplot 생성 유무
* patience : 더이상 monitor 값의 개선이 없는 경우, 학습을 Early Stopping하는 기준

[참고] Patience 가 10인 경우 Early Stopping 결과 예시

텍스트이(가) 표시된 사진

자동 생성된 설명

* + - 1. Fit 설정 및 학습 시작
* train\_data : 학습을 위한 데이터
* validation\_data : 학습 중, 매 epoch( 1 Iteration\_threshold) 마다 검증하기 위한 데이터
* epochs : 학습 횟수
* full\_shuffle\_flag : 모든 입력 데이터를 shuffle 해서 학습을 할것인지(default : True), Bagging 방식으로 학습할 것인지(False) 선택

[참고] Bagging (Bootstrap Aggregating) : 학습 데이터의 다양성을 위해 무작위로 복원 추출하여 데이터를 추출하는 방식

* + - 1. 학습 모델 저장
* esoinn\_{모델 버전}\_{epoch 횟수}.pickle 파일명으로 binary 형태 객체 저장
  1. Prediction model precess
     1. 필수 페케지 import (3.1.1 참고)

* + 1. 데이터 생성 (3.1.2 참고)
    2. 예측 데이터 생성

텍스트, 화면, 스크린샷이(가) 표시된 사진

자동 생성된 설명

* + 1. Prediction

텍스트이(가) 표시된 사진

자동 생성된 설명

* + - 1. 학습된 모델 load
* 모델 버전과 epoch 값을 받아 학습된 모델을 load 한다.
  + - 1. 예측 실행
* 데이터를 입력 받아, 공격 데이터와 일반 데이터의 index를 list로 반환한다.
  + - 1. 모델 평가
* 공격과 일반(‘normal’)에 대한 입력 데이터의 라벨이 있는 경우, 예측 모델의 accuracy, precison, recall, f1\_score를 출력한다.l

# Trouble shooting

# reference

1. Furao, S., et al. (2007). "An enhanced self-organizing incremental neural network for online unsupervised learning." Neural networks 20(8): 893-903
2. Furao, S. and O. Hasegawa (2006). "An incremental network for on-line unsupervised classification and topology learning." Neural networks 19(1): 90-106
3. ESOINN github https://github.com/huagc/ESoinn.git