**NIA 프로젝트 주간보고서(11월 2주차)**

|  |  |  |  |
| --- | --- | --- | --- |
| **팀원** | 김경수, 이세림, 이지상 | | |
| **작성자** | 김경수, 이지상, 이세림 | **작성일** | 2021.11.24(수) |

1. **작성자**
2. **주요 진행사항**

|  |  |
| --- | --- |
| **이슈사항** |  |

1. **세부 진행사항**

|  |  |
| --- | --- |
| **요 일** | **업무내용** |
| **11월 19일(금)** | * **모델 데이터 구조 분석(1)**      * + ambient\_temperature\_system\_failure.csv : 사무실 설정의 주변 온도   + cpu\_utilzation\_asg\_misconfiguration.csv : 주어진 클러스터의 평균 CPU 사용량   + ec2\_request\_latency\_system\_failure.csv : 서버의 CPU 사용량 데이터   + machine\_temperature\_system\_failure.csv : 대형 산업용 기계의 내부 구성 요소에 대한 온도 센서 데이터   + nyc\_taxi.csv : NYC 택시 승객 수   + rogue\_agent\_key\_hold.csv : 컴퓨터의 여러 사용자에 대해 키가 유지되는 타이밍   + rogue\_agent\_key\_updown.csv : 컴퓨터의 여러 사용자에 대한 키 입력 타이밍   + 위는 VAE-LSTM 모델을 학습 및 평가할 때 사용한 데이터셋(csv파일)     …     * + 내부 파일은 timestamp와 value로 이루어져 있지만 이것이 무엇을 의미하는지 아직 확실하지 않음 |
| **11월 20일(토)** | * **모델 데이터 구조 분석(2)**      * + ambient\_temperature\_system\_failure.csv     - 1시간 주기로 측정한 온도(화씨)      * + cpu\_utilzation\_asg\_misconfiguration.csv     - 5분 단위로 측정한 클러스터의 CPU 사용량     - 최대 100, 실수값      * + ec2\_request\_latency\_system\_failure.csv     - 5분 단위로 측정한 서버의 CPU 사용량     - 최대 100, 실수값      * + machine\_temperature\_system\_failure.csv     - 5분 단위로 측정한 대형 산업용 기계의 온도 센서 데이터      * + nyc\_taxi.csv     - 30분 단위로 측정한 NYC 택시의 총 이용 승객 수      * + rogue\_agent\_key\_hold.csv     - 컴퓨터 여러 사용자에 대하여 키가 유지되는 타이밍     - 다수에 의해 변경되므로 매우 작은 값이 나올거라 예상      * + rogue\_agent\_key\_updown.csv     - 컴퓨터의 여러 사용자에 대한 키 입력 타이밍     - 이 데이터와 rogue\_agent\_key\_hold.csv에 대한 정보가 확실하지 않음 * 해당 데이터는 무언가의 원인에 대한 정보를 담고 있지만 E2ON과 같이 특정 행동에 대한 이상행동 정보를 담고 있는 것은 아님 |
| **11월 21일(일)** | * **모델 구조 분석(1)**   + 깃에 나와있는 데이터는 일반적으로 시계열 데이터로 시간대별로 해당하는 값이 들어 있음   + 해당 깃허브에는 NAB(numenta Anomaly Benchmark)를 이용하여 이상감지를 평가 하는걸로 추정 하지만 여기서 이상 추정은 시간대별로 값을 점수로 계산한후 그래프로 나타낸다음 그 점수가 도드라지거나 완전히 감소되거나 증간한 횟수나 값으로 평가하는 것으로 보임 |
| **11월 22일(월)** | * **모델 데이터 구조 분석(2)**   + 출력결과는 F1 score와 가장 높은 정확도의 threshold와 해당 threshold에 따른 정밀도, 평균 정밀도가 출력된다. 기본 출력과 더불어 augmented detection 결과가 출력 되는데 해당 결과는 기존 결과에서 정확도를 높이기 위해서 한번 더 sort를 해준 것으로 판단되나 확실하지 않음 |
| **11월 23일(화)** | * **이상행동 모델 재탐색**   + VAE-LSTM 모델이 E2ON에서 제공된 데이터로는 학습이 가능한 형태로 변경이 불가능하다고 판돤되어 다른 모델들을 재탐색   + 1. C3D와 I3D를 사용한 Sultani 모델      * + 2. RNN 기반 시계열데이터 이상 탐지 RNN-Time-series 모델   arch   * + 여러 모델들을 찾아봤으나 학습데이터 구조가 정확히 공개되어 있는 모델이 없어 더 탐색해야함 |

1. **기타 보고사항**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **차주  진행계획** | * 11월 5주차 진행계획  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **구분** | **11월 26일** | **11월 27일** | **11월 28일** | **11월 29일** | **11월 30일** | | **연구**  **내용** | 모델 재탐색  (2) | 모델 재탐색  (3) | 학습데이터  전처리 | 모델  학습 및  평가 | 주간보고서 작성 및  발표 자료 준비 | | **산출물** | - | - | 학습데이터 | 평가 결과 | 주간 보고서및 발표자료 | |
| **특이사항** |  |