**NIA 프로젝트 주간보고서(12월 1주차)**

|  |  |  |  |
| --- | --- | --- | --- |
| **팀원** | 김경수, 이세림, 이지상 | | |
| **작성자** | 김경수, 이지상, 이세림 | **작성일** | 2021.12.01(수) |

1. **작성자**
2. **주요 진행사항**

|  |  |
| --- | --- |
| **이슈사항** | * **MPED-RNN 모델과 VAE-LSTM 모델 사용** * VAE-LSTM 모델을 사용하기 위해선 단일 값의 value가 필요한데, 좌표값이 value값이 되기에는 무리가 있어보여 MPED-RNN 모델에서의 y\_hat을 value로 사용하기 위해 MPED-RNN 모델 사용 * 성공 시 value에 좌표를 넣어보는 등의 또다른 방법 시도 예정 |

1. **세부 진행사항**

|  |  |
| --- | --- |
| **요 일** | **업무내용** |
| **11월 26일(금)** | * **모델 확인**   + 주간미팅 때 받은 피드백 대로 모델을 찾는 것이 아니라 input의 형태를 변경하여 진행해보기 위하여 지금까지 찾아봤던 모델 중 가장 적합한 모델을 찾기 위해 재확인 진행     모델에 대한 의견   * + VAE-LSTM 모델이 현재 가장 적합하다고 판단되어 VAE-LSTM으로 진행 |
| **11월 27일(토)** | * **VAE-LSTM 모델을 통한 검증(1)**     MPED-RNN 모델의 중간 결과(Y\_Hat)    VAE-LSTM의 훈련 데이터(좌)와 Y\_Hat을 통해 제작한 데이터(우)   * MPED-RNN 모델에서 나온 중간 결과인 Y\_Hat을 저장하여 기존의 훈련 데이터와 유사한 구조로 제작함 |
| **11월 28일(일)** | * **VAE-LSTM 모델을 위한 데이터 분석**      * 현 모델에서 사용하는 데이터셋은 csv, npz 확장자 파일이 존재하며 데이터셋에 대한 전처리 코드가 존재함      * 좌측은 인덱스명, 우측은 인덱스에 대한 value값 * t는 csv 각 값의 순서이며, t\_unit은 t가 1씩 늘어날 때마다 증가하는 시간 단위 * readings는 value이며, idx\_anomaly는 이상 t로 추정(위 사진으로는 두군데) * idx\_split은 train과 test가 나뉘는 구간으로 추정 * training과 test는 각각 t\_train과 t\_test의 개수만큼 존재하며 각 time에 따른 값이 존재 * train\_m은 학습 평균, train\_std는 편차를 의미함 * idx\_anomaly\_test는 test에서의 이상부분 t로 추정        * 전처리 코드의 실행 시 출력되는 결과 * 파란 선은 idx\_split, 빨간 선은 idx\_anomaly * 각 지표의 최상점과 최하점을 idx\_anomaly로 판단하는 것으로 추정 |
| **11월 29일(월)** | * **VAE-LSTM 모델을 위한 E2ON 데이터 수정(1)** * VAE-LSTM은 학습과 평가를 한 번에 하는 모델이기 때문에 학습 후 만들어지는 파일 npz의 구조를 맞춰주기 위한 코드 작성과 데이터셋 전처리가 필요 * 전처리를 하기 위해선 총 12개의 정보가 필요      * t는 1부터 데이터의 개수까지의 순서로 frame의 개수만큼 np.arange * t\_unit은 증가하는 단위로 ‘frame’ * readings는 csv파일에선 value값으로 되어 있으나 1차원 리스트의 형태로 되어 있어 x, y 좌표를 넣는 것은 불가능하기에 MPED-RNN 모델을 통해 얻는 y\_hat 값을 입력값으로 할당 🡺 실수값을 할당한 후 성공시 좌표로 시도 예정      * idx\_split은 train과 test로 나뉘는 구간으로 볼 수 있기 때문에 8:1:1 비율을 유지하기 위해 우선 나뉘는 구간을 frames\*0.8로 할당      * train\_m과 train\_std는 각각 평균과 편차이기 때문에 함수를 사용하여 값 할당 * t\_train과 t\_test도 각각 학습과 평가에 대한 데이터셋 개수에 대한 순서 * 하지만 현재 값에 대해 값이 만들어지는 과정이 불분명한 idx\_anomaly, training, test, anomaly\_test에 대하여 정보를 더 찾아볼 필요가 있음 |
| **11월 30일(화)** | * **VAE-LSTM 모델을 위한 E2ON 데이터 수정(2)** * 이상행동 frame을 선택하는 것은 주석파일 혹은 frame level masks 파일에서 가져와 frame을 선택하도록 판단      * 기존 idx\_anomaly는 이상 현상의 frame 시간대를 미리 정해두고 파일 제작 * E2ON의 경우 이상행동의 frame이 연속으로 이어지기 때문에 해당하는 모든 frame을 anomalies에 넣어 학습 후 평가가 이뤄질 수 있도록 변환 예정      * 따라서 idx\_anomaly는 위와 같이 표현이 가능함 * 다만 이상현상에 대한 연속적인 정보기 때문에 MPED-RNN에서 사용한 frame\_level\_mask파일(혹은 E2ON json 파일)을 가지고 값을 할당할 예정      * training과 test는 (readings – train\_m) / train\_std에 대한 값으로 frames\*0.8에 대한 값 전후로 계산을 하여 넣어주는 과정을 반복 * anomaly\_test는 test에 대한 데이터 중 이상이 있는 t를 말하는 1차원 리스트 * (①)MPED-RNN 모델에서 y\_hat을 얻어 (②)전처리를 통해 만든 npz를 (③)VAE-LSTM에 사용해 평가하는 과정을 한 번에 처리할 수 있도록 하는 중에 있음 |

1. **기타 보고사항**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **차주  진행계획** | * 12월 2주차 진행계획  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **구분** | **12월 3일** | **12월 4일** | **12월 5일** | **12월 6일** | **12월 7일** | | **연구**  **내용** | 전처리와 학습 및 평가 과정 처리 | E2ON 데이터셋 csv파일 생성 | 학습 및 평가(1) | 검토, 학습 및 평가(2) | 주간보고서 작성 및  발표 자료 준비 | | **산출물** | 과정이 간략화 된 코드, 깃 업로드 | - |  |  | 주간 보고서및 발표자료 | |
| **특이사항** |  |