**NIA 프로젝트 주간보고서(11월 2주차)**

|  |  |  |  |
| --- | --- | --- | --- |
| **팀원** | 김경수, 이세림, 이지상 | | |
| **작성자** | 김경수, 이지상, 이세림 | **작성일** | 2021.11.17(수) |

1. **작성자**
2. **주요 진행사항**

|  |  |
| --- | --- |
| **이슈사항** | * **VAE-LSTM 학습** * 평가 방법 강구와 학습 데이터 구조 분석 필요 |

1. **세부 진행사항**

|  |  |
| --- | --- |
| **요 일** | **업무내용** |
| **11월 12일(금)** | * **파이썬 tensorflow 스터디 및 모델 탐구**   + 딥러닝/머신러닝에 필요한 라이브러리 중 VAE-LSTM 모델에 필요한 tensorflow 라이브러리 집중 스터디   + 모델에서 사용하는 대표적인 함수를 위주로 스터디 후 이후 다른 tensorflow 함수와 keras 등 다양한 라이브러리 스터디 예정   + 해당 모델은 tensorflow v.1을 사용   + 추가적으로 matplotlib와 가장 기초적인 numpy에 대해서도 사용하는 tensorflow 함수 스터디를 하면서 같이 스터디  |  |  | | --- | --- | | train.py | 모델 학습을 위한 py파일 | | data\_loader.py | 데이터셋을 불러오는 py파일 | | models.py | 모델 선언 py파일, VAE 모델과 LSTM 모델이 있음 | | trainers.py | VAE 모델 학습을 위한 기능 py파일 | | utils.py | 데이터셋 처리 전처리 및 args 사용을 위한 py파일 | | base.py |  |  * + 모델의 구현은 잘 되어있다고 볼 수 있으나 모델에 대한 설명이 깊지 않아 계속해서 살펴볼 필요가 있음 |
| **11월 13일(토)** | * **VAE-LSTM 모델 확인**   + 기본 데이터셋은 csv 파일과 이를 이용할 수 있는 npz파일로 구성     < 좌 npz파일, 우 csv-files >   * + 모델 훈련에 대한 코드는 이미 정의되어 있고, 학습 후 훈련된 모델을 사용해 이상 감지가 가능   + 기본적으로 버전은 기존보다 다운된 것을 사용해야 함      * + codes/train.py 학습 후 codes/NAB-anomaly-detection.ipynb을 이용해 이상 감지   + NAB-anomaly-detection.ipynb는 NAB-config.json 파일이 필요하며 그것은 다음과 같음 |
| **11월 14일(일)** | * **VAE-LSTM 학습 데이터 구조 분석**   + Timestamp와 value로 이루어진 csv파일     …     * + nyc\_taxi.csv의 경우 30분 간격(Timestamp)과 정수형태의 value값으로 총 10320개의 정보가 나타나 있음        * + 그 외의 파일 또한 한시간 또는 5분 간격으로 정보가 저장되어 있으며, 나머지 파일은 실수형태의 value값임   + 각 값에 대한 정확한 의미는 파악중 |
| **11월 15일(월)** | * **VAE-LSTM 학습 (1)**        * + 모델에서 요구하는 텐서플로우 버전인 1.15.2로 실행 시 2.6 이상이 필요하다고 오류      * + 텐서플로우 버전을 2.6.0으로 업그레이드 후 실행 시 텐서플로우 1.x 버전에서만 지원되는 Session 부분에서 오류 발생   + 학습파일로 timestamp와 value 로 이루어진 파일을 사용하는데 해당 value 값이 무엇을 의미하는지 확인 필요 |
| **11월 16일(화)** | * **VAE-LSTM 학습 (2)**   + 완전히 새로운 가상환경을 만들어 라이브러리 설치        * + 학습이 되는 것을 확인하였으나 평가가 가능한 코드가 없음.   + 논문에 나와 있는 F1 SCORE를 출력 하기 위해 코드 수정 필요 |

1. **기타 보고사항**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **차주  진행계획** | * 11월 4주차 진행계획  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **구분** | **11월 19일** | **11월 20일** | **11월 21일** | **11월 22일** | **11월 23일** | | **연구**  **내용** | 모델 데이터 구조 분석(1) | 데이터 구조에 따른 E2ON 데이터셋 전처리(1) | 데이터 구조에 따른 E2ON 데이터셋 전처리(2) | 전처리된 데이터셋을 이용한 학습과 평가 | 주간보고서 작성 및  발표 자료 준비 | | **산출물** | - | - | 전처리된 E2ON 데이터셋 | 평가 결과 | 주간 보고서및 발표자료 | |
| **특이사항** |  |