**NIA 프로젝트 주간보고서(9월 1주차)**

|  |  |  |  |
| --- | --- | --- | --- |
| **팀원** | 김경수, 이세림, 이지상 | | |
| **작성자** | 이세림 | **작성일** | 2021.09.8(수) |

1. **작성자**
2. **주요 진행사항**

|  |  |
| --- | --- |
| **이슈사항** | * **AI-Hub 데이터셋 중 학습, 평가가 가능한 영상 증가** * 기존에 유지하던 지수 형태의 csv 파일 구조로는 108개 영상 중에 57개가 학습이나 평가 시 value 에러가 나왔으나 정수 형태로 바꾼 후 108개 중 6개만 value 에러가 발생 * 기존엔 51개 영상으로 학습, 평가를 진행하였으나 현재 102개 영상을 78:24로 나누어 학습, 평가 * **출력 방식 변경의 문제** * 현재 코드 내에 정의된 predict() 함수나 sklearn에서 제공되는 roc\_auc\_score() 함수를 통해 AUROC, AUPR을 출력으로 사용하고 있음 * accuracy 출력을 위해 함수 제작과 코드 변경 중에 있음 |

1. **세부 진행사항**

|  |  |
| --- | --- |
| **요 일** | **업무내용** |
| **9월 3일(금)** | * **AI-Hub 데이터셋 value 오류 해결을 위한 회의(1)**  |  | | --- | | ValueError: Input contains infinity or a value too large for dtype('float32'). |  * + 학습, 평가까지 진행은 하였으나 value 오류가 발생하여 제공받은 영상 108개 중에 57개 밖에 사용하지 못하고 있음   + 오류가 발생하는 csv 파일을 특정하고 각 행을 지워가면서 오류가 발생하는 행을 특정하였으나 오류가 발생하지 않는 다른 행과의 차이점을 발견하지 못함 |
| **9월 4일(토)** | * **AI-Hub 데이터셋 value 오류 해결을 위한 회의(2)**   + 기존에 사용하고 있는 csv 파일은 지수 형태로 사용하였으나 csv 파일 구조를 정수로 바꾸어 학습   + 오류가 발생하는 파일이 57개에서 6개로 대폭 감소함   기존 csv 파일  정수 형태로 바꾼 csv 파일   * + 정수로 구조를 바꿨음에도 계속 오류가 발생하는 파일들에 대한 대책 강구 필요     - 텍스트이(가) 표시된 사진        자동 생성된 설명소수점 반올림의 자리를 한 자리씩 미뤄가며 오류를 해결해볼 예정   기존에 학습, 평가되지 않는 영상    정수형으로 구조를 바꾼 뒤 학습, 평가되지 않는 영상 |
| **9월 5일(일)** | * **AI-Hub 데이터셋 재학습 및 재평가**   + 제공받은 108개 영상 중 계속 오류가 발생하는 6개를 제외한 102개 영상으로 학습 평가를 진행   + 텍스트이(가) 표시된 사진      자동 생성된 설명텍스트이(가) 표시된 사진      자동 생성된 설명102개 영상을 78:24로 나누어서 학습:평가를 진행 하였으며 평가 데이터는 이상행동 종류 12개를 각각 2개씩 사용   기존 결과(좌) csv 파일 변경 후 결과(우)   * + 기존 결과와 큰 차이를 보이지는 않음 |
| **9월 6일(월)** | * **모델 출력 결과 변경**   + 평가 방식에 관련된 py파일은 여러가지가 존재     - evaluate.py, tbad/eval.py, tbad/combined\_model/evaluate.py 등   + predict함수는 모델 평가가 아닌 모델을 사용한다는 것으로 보는 것이 맞으며 위 함수는 평가보다는 정확도를 판별하는 것에 도움을 줌   + evaluate를 실행하면 predict함수가 동작을 하며 중간에 아래와 같은 코드를 실해함   + 원하는 형태의 결과는 evaluate 함수를 사용했을 때 리턴되는 loss, acc값      * + evaluate함수를 사용하지 않고 값을 얻으려면 predict의 결과값을 가지고 loss와 acc을 구해야 하기 때문에 model.evaluate()함수를 사용하기 위한 방법(구현) 논의 |
| **9월 7일(화)** | * **모델 출력 결과**   + predict()함수를 응용해 evaluate()함수 작성      * + 시험삼아 학습 후 evaluate()함수를 사용해보았지만 Syntax Error가 뜨며 실행되지 않음   + evaluate()함수를 사용하기 위한 코드를 작성하기 어려워 현재 학습을 하면서 후반에 출력을 하게 할 지, 따로 평가를 위한 코드를 작성할지 현재 의논 중 |

1. **기타 보고사항**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **차주  진행계획** | * 9월 3주차 진행계획  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **구분** | **9월 10일** | **9월 11일** | **9월 12일** | **9월 13일** | **9월 14일** | | **연구**  **내용** | AI-Hub 데이터셋 오류 보완 | evaluate 함수 사용을 위한 구문 탐색 | evaluate 함수 사용을 위한 구문 보완(1) | evaluate 함수 사용을 위한 구문 보완(2) | evaluate 함수를 사용한 모델 평가 | | **산출물** | AI-Hub 데이터셋 csv파일 |  |  | 평가에 맞는 train.py 또는 evaluate.py 파일 | 프레임 또는 epoch별 모델 평가 | |
| **특이사항** | * **여전히 발생하는 일부 파일의 value 오류**   + Point 값을 scale 하는 중에 발생하는 것으로 추정 * **평가 지표 수정의 어려움**   + 평가에 대한 정보를 출력하는 기존 함수의 개수가 많으며 각 함수마다 고유의 형태를 지닌 것이 아닌 함수 간의 연결성으로 수정의 어려움을 겪고 있음 |