**적합도 검사서**

## **1. 데이터**

**학습데이터**

무신사 스탠다드의 ‘레이어드 크루 넥 반팔 티셔츠\_일반 기장 2팩’

상품의 1~4월 리뷰데이터, 데이터 갯수 607개

**테스트 데이터** :

무신사 스탠다드의 ‘레이어드 크루 넥 반팔 티셔츠\_일반 기장 2팩’

상품의 5~7월 리뷰데이터, 데이터 갯수 197개

**이벤트 로그 추출과정**

뽀로로 개체명 추출(Named Entity Recognition)

카카오 브레인의 뽀로로 라이브러리를 사용하여 개체명 추출

| review\_ID | date | ner |
| --- | --- | --- |
| review\_03283 | 2024-01-31 02:44:42 | [('티셔츠', 'CIVILIZATION'), ('티셔츠','CIVILIZATION')] |
| review\_02512 | 24-01-30 20:27:01 | [('반팔티', 'CIVILIZATION'), ('무텐다드', 'ORGANIZATION'), ('올해', 'DATE'), ('아이폰', 'ARTIFACT'), ('반팔티', 'CIVILIZATION'), ('1만원대', 'QUANTITY'), ('무텐다드', 'ORGANIZATION')] |

**이벤트 형태로 가공**

하나의 리뷰안에 들어있는 여러개의 개체명을 타임스탬프를 이용해 개체를 구분할 수 있도록 가공

[['review\_03283', '2024-01-31 21:20:47', 'CIVILIZATION'],

['review\_03283', '2024-01-31 21:20:48', 'CIVILIZATION'],

['review\_02512', '2024-01-30 10:48:42', 'CIVILIZATION'],

['review\_02512', '2024-01-30 10:48:43', 'ORGANIZATION'],

['review\_02512', '2024-01-30 10:48:44', 'DATE'],

['review\_02512', '2024-01-30 10:48:45', 'ARTIFACT'],

['review\_02512', '2024-01-30 10:48:46', 'CIVILIZATION'],

['review\_02512', '2024-01-30 10:48:47', 'QUANTITY'],

['review\_02512', '2024-01-30 10:48:48', 'ORGANIZATION']

## **2.적용한 평가 방식**

**Token-Based-Replay**

모델과 프로세스가 얼마나 일치하는지 확인하는 적합성 검사 알고리즘

학습데이터로 도출한 프로세스가 테스트 데이터와 얼마나 일치하는 지 확인

전체 로그에서 모델에 잘 맞는 추적비율 (\*\*perc\_fit\_traces): 97.29,

평균 추적 적합도(\*\*average\_trace\_fitness): 0.99,

전체 로그의 적합도(\*\*log\_fitness): 0.99,

로그 내 잘 맞는 추적 비율(\*\*percentage\_of\_fitting\_traces): 97.29

**학습데이터로 도출한 ‘프로세스’와 테스트 데이터의 일치율은**

**약 97.29%인 것을 확인 할 수 있다.**