

2022 NIKL Artificial Intelligence Language Capability Evaluation Competition Corpus_Aspect-Based Sentiment Analysis

(version 1.0)

- **Title:** 2022 NIKL Artificial Intelligence Language Capability Evaluation Competition Corpus – Aspect-Based Sentiment Analysis

- **Release Date**

- (version 1.0) 16 September 2022

- **Data Type:** text

- **Description**

- Outline

Aspect-Based Sentiment Analysis refers to categorizing sentimental qualities of an entity in language use.

- Corpus size: 7,917 sentences in total

	train	validation	test
number of sentences	3,001	2,794	2,127

※ The annotations of the test dataset are not released.

- File Format and Encoding: JSONL(UTF-8)

- Number and Size of Files: 3 files, 1.61MB in total

- **Citation(s):**

(Korean) 국립국어원(2022), 2022 국립국어원 인공 지능 언어 능력 평가 대회 말뭉치-속성 기반 감성 분석(버전 1.0) URL: <https://kli.korean.go.kr/corpus>

(English) The National Institute of Korean Language (2022), 2022 The National Institute of Korean Language Artificial Intelligence Language Capability Evaluation Competition Corpus – Aspect-Based Sentiment Analysis (version 1.0) URL: <https://kli.korean.go.kr/corpus>

• Examples

```
{
  "id": "nikluge-sa-2022-train-00001",
  "sentence_form": "둘쨋날은 미친듯이 밋아봤더니 기어가 헛돌면서  
탁탁 소리가 나서 경악.",
  "annotation": [
    ["본품#품질", ["기어", 16, 18], "negative"]
  ]
}
{
  "id": "nikluge-sa-2022-train-00002",
  "sentence_form": "이거 뭐 뺑꾸를 준 거 아냐 불안하고, 거금 투자  
한 게 왜 이래.. 싫어서 정이 확 떨어졌는데 산 곳 가져가서 확인하니 기어 텐션 문제라고 고장 아니래.",
  "annotation": [
    ["본품#품질", ["기어 텐션", 67, 72], "negative"]
  ]
}
{
  "id": "nikluge-sa-2022-train-00003",
  "sentence_form": "간사하게도 그 이후에는 라이딩이 아주 즐거워져  
서 만족스럽게 탔다.",
  "annotation": [
    ["제품 전체#일반", [null, 0, 0], "positive"]
  ]
}
{
  "id": "nikluge-sa-2022-train-00004",
  "sentence_form": "삭이 없는 모델이라 일반 도로에서 타면 노면의  
진동 때문에 손목이 덜덜덜 떨리고 이가 부딪칠 지경인데 이마저도 며칠 타면서 익숙해지니 신경쓰이지 않게  
됐다.",
  "annotation": [
    ["제품 전체#일반", ["삭이 없는 모델", 0, 8], "neutral"]
  ]
}
{
  "id": "nikluge-sa-2022-train-00005",
  "sentence_form": "안장도 딱딱해서 엉덩이가 아팠는데 무시하고 타  
고 있다.",
  "annotation": [
    ["본품#일반", ["안장", 0, 2], "negative"]
  ]
}
```