

## 국립국어원 속성 기반 감성 분석 말뭉치 2021

(버전 1.0)

- 자료명: 국립국어원 속성 기반 감성 분석 말뭉치 2021
- 공개일
  - (버전 1.0) 2022. 4. 1.
- 자료 유형: 텍스트
- 관련 사업: 2021년 말뭉치 감성 분석 및 연구(2021)
- 자료 설명
  - ※ 자세한 내용은 국립국어원 누리집 > 자료 > 연구·조사 자료 > '2021년 말뭉치 감성 분석 및 연구' 사업 결과보고서 참조
- 내용
  - 국립국어원 감성 분석 말뭉치 2020(1.0)과 동일한 문서에 속성 기반 감성 정보 부착
  - “SemEval-2016 Task-5: Aspect-Based Sentiment Analysis”의 방법론을 따라 말뭉치를 구축하였으며, 분석 대상 문서의 분량구성을 고려하여 개체#속성 쌍을 정의

<표 1> 영화 도메인의 개체#속성 쌍

	일반	구성력	연기력	인지도	연출력
영화	0			0	0
시나리오	0	0			
배우	0		0	0	
캐릭터	0				0
감독	0			0	0
음향·음악	0			0	0
영상·시각 요소	0				0

<표 2> 제품 도메인의 개체#속성 쌍

	일반	가격	디자인	품질	편의성	다양성	인지도
제품 전체	0	0	0	0	0		0
본품	0		0	0	0	0	
패키지·구성품	0		0	0	0	0	
브랜드	0	0	0	0			0

<표 3> 여행 도메인의 개체#속성 쌍

	일반	가격	편의성	심미성	접근성	청결도	인지도
장소 전체	0			0	0	0	0
시설	0	0	0	0		0	
상품	0	0					0
서비스	0						

· 분량

- 이 말뭉치는 “2022년 국립국어원 인공 지능 언어 능력 평가” 대회 과제로 활용 검토중이므로 “국립국어원 감성 분석 말뭉치 2020(1.0)” 대상 문서 중 50% 공개(1,026개 문서)

<표 4> ‘2021년 말뭉치 감성 분석 및 연구’ 사업의 말뭉치 구축 결과\*

분석한 문장 수(문서 수)

대상 자료	영화	제품			여행			합계
		화장품 세정제	전자 기기	제품 기타	숙소	식당	관광지	
감성 분석 말뭉치 2020(1.0)	3,273 (379)	2,939 (775)	700 (123)	2,559 (640)	265 (27)	266 (35)	712 (74)	10,714 (2,053)
추가 구축**	1,094 (69)	1,545 (67)	1,002 (68)	-	-	-	-	3,641 (204)
합계	4,367 (448)	3,941 (842)	2,245 (191)	2,559 (640)	265 (27)	266 (35)	712 (74)	14,355 (2,257)

\* 공개 자료와 비공개 자료 모두 포함한 총 구축 결과

\*\* 자료 활용성 제고를 위해 본 사업에서 추가 구축한 자료 규모

· 파일 형식: JSON(UTF-8 인코딩)

· 파일 수 및 크기: 파일 2개, 총 1.30MB(ZIP 압축)

· 인용: 국립국어원(2022). 국립국어원 속성 기반 감성 분석 말뭉치 2021 (버전 1.0).

URL: <https://corpus.korean.go.kr>

· 파일 명명 규칙

자리	1	2	3	4	5	6	7	8	9	10	11	12	13	14
속성	매체	장르	주석	단계	구축 연도	문서 / 문장	일련번호(7자리)							
정의값	E: 웹	X: 추출	SA: 감성		21: 2021년	1: 문서 2: 문장	0000001 ~ 9999999 (일곱 자리 일련번호)							
※ 예시: EXSA2112203180.json 2021년도에 구축한 문서 단위 감성 분석 말뭉치 파일														
※ 예시: EXSA2122203180.json 2021년도에 구축한 문장 단위 감성 분석 말뭉치 파일														

· 예시

· sentence\_level

```
{
  "document_id": "ERRW1903002756.23",
  "domain": "영화",
  "is_test": "N",
  "metadata": {
    "title": "놈놈놈",
    "author": "김마리짱",
    "publisher": "egloos",
    "date": "20080718",
    "url": "http://kimmari02.egloos.com/1876418"
  },
  "sentence": [
    {
      "sentence_id": 1,
      "sentence_form": "영화도 없고 그냥 화제작이라는 건 늘 개봉 첫날 구미가 바싹 당기는 법이니까 다녀왔다.",
      "opinions": [
        {
          "opinion_id": 1,
          "category": "영화#인지도",
          "opinion_target": "화제작",
          "begin": 46,
          "end": 49,
          "polarity": "neutral"
        }
      ]
    },
    {
      "sentence_id": 2,
      "sentence_form": "할아버지까지 넘쳐나더군요. - -:",
      "opinions": []
    },
    {
      "sentence_id": 3,
      "sentence_form": "할아버지까지 넘쳐나더군요. - -:",
      "opinions": []
    },
    {
      "sentence_id": 4,
      "sentence_form": "같다.",
      "opinions": [
        {
          "opinion_id": 1,
          "category": "영화#인지도",
          "opinion_target": "화제작",
          "begin": 46,
          "end": 49,
          "polarity": "neutral"
        }
      ]
    }
  ]
}
```

```

        "opinion_id": 1,
        "category": "영화#인지도",
        "opinion_target": "Null",
        "begin": 0,
        "end": 0,
        "polarity": "positive"
      }
    ]
  },
  {
    "sentence_id": 5,
    "sentence_form": "애초에 기대를 반쯤 접고 봐서 그런지 나쁘지는 않았는데 사실은
    좀 흐름.. 했다.",
    "opinions": [
      {
        "opinion_id": 1,
        "category": "영화#일반",
        "opinion_target": "Null",
        "begin": 0,
        "end": 0,
        "polarity": "neutral"
      }
    ]
  },
  {
    "sentence_id": 6,
    "sentence_form": "스토리 라인이 너무 빈약하더군.",
    "opinions": [
      {
        "opinion_id": 1,
        "category": "시나리오#구성력",
        "opinion_target": "스토리 라인",
        "begin": 0,
        "end": 6,
        "polarity": "negative"
      }
    ]
  },
  {
    "sentence_id": 7,
    "sentence_form": "물론 후반부의 추격씬이랄까 싸우는 광경은 꽤 잘 찍었다.",
    "opinions": [
      {
        "opinion_id": 1,
        "category": "영상/시각 요소#연출력",
        "opinion_target": "싸우는 광경",
        "begin": 15,
        "end": 21,
        "polarity": "positive"
      }
    ]
  },
  {
    "sentence_id": 8,
    "sentence_form": "정말 웨스턴 무비 같다.",
    "opinions": []
  },
  {
    "sentence_id": 9,
    "sentence_form": "이런 장르가 한국영화에는 없었기 때문에 더욱 실감나기도 하고..",
    "opinions": [
      {
        "opinion_id": 1,

```

```

        "category": "영화#연출력",
        "opinion target": "Null",
        "begin": 0,
        "end": 0,
        "polarity": "positive"
    }
}
},
{
    "sentence_id": 10,
    "sentence_form": "송강호의 `이상한 놈`이 주인공 같은데 평소 캐릭터를 십분 살려서
    곳곳에 유머를 배치하려고 애쓴 것도 보이고 물론 나도 몇 번 웃긴 했는데 그냥 피식 정도..",
    "opinions": [
        {
            "opinion_id": 1,
            "category": "캐릭터#일반",
            "opinion target": "송강호의 `이상한 놈",
            "begin": 0,
            "end": 11,
            "polarity": "neutral"
        }
    ]
},
{
    "sentence_id": 11,
    "sentence_form": "어떤 건 너무 의도적으로 웃기려고 한 대사라 전혀 우습지가
    않았다.",
    "opinions": [
        {
            "opinion_id": 1,
            "category": "시나리오#일반",
            "opinion target": "대사",
            "begin": 21,
            "end": 23,
            "polarity": "negative"
        }
    ]
},
{
    "sentence_id": 12,
    "sentence_form": "강철중 같은 대사발을 기대하면 필시 실망할 듯..",
    "opinions": [
        {
            "opinion_id": 1,
            "category": "OUT OF SCOPE",
            "opinion target": "Null",
            "begin": 0,
            "end": 0,
            "polarity": "positive"
        }
    ]
},
{
    "sentence_id": 13,
    "sentence_form": "그래서 재미없다는 건 아니고 액션이 관촬았기 때문에 중박
    정도라고 생각한다.",
    "opinions": [
        {
            "opinion_id": 1,
            "category": "영상/시각 요소#연출력",
            "opinion target": "액션",
            "begin": 16,
            "end": 18,

```

```

        "polarity": "positive"
      },
      {
        "opinion_id": 2,
        "category": "영화#일반",
        "opinion target": "Null",
        "begin": 0,
        "end": 0,
        "polarity": "neutral"
      }
    ]
  },
  {
    "sentence_id": 14,
    "sentence_form": "이병헌과 정우성의 캐릭터는 그야말로 딱 중간 정도..",
    "opinions": [
      {
        "opinion_id": 1,
        "category": "캐릭터#일반",
        "opinion target": "이병헌과 정우성의 캐릭터",
        "begin": 0,
        "end": 13,
        "polarity": "neutral"
      }
    ]
  },
  {
    "sentence_id": 15,
    "sentence_form": "확실히 송강호가 왜 맨 첫번째로 이름이 나왔는지 알 거 같았다.",
    "opinions": [
      {
        "opinion_id": 1,
        "category": "배우#일반",
        "opinion target": "송강호",
        "begin": 4,
        "end": 7,
        "polarity": "positive"
      }
    ]
  },
  {
    "sentence_id": 16,
    "sentence_form": "분명 나이순은 아닌 것 같다. -_-;",
    "opinions": []
  },
  {
    "sentence_id": 17,
    "sentence_form": "그 중간이라는 게 아쉬운 거지..",
    "opinions": [
      {
        "opinion_id": 1,
        "category": "캐릭터#일반",
        "opinion target": "Null",
        "begin": 0,
        "end": 0,
        "polarity": "negative"
      }
    ]
  },
  {
    "sentence_id": 18,
    "sentence_form": "칭찬하고 싶은 건 영화 음악.",
    "opinions": [

```

```

        {
            "opinion_id": 1,
            "category": "음향/음악#일반",
            "opinion_target": "영화 음악",
            "begin": 10,
            "end": 15,
            "polarity": "positive"
        }
    ],
    },
    {
        "sentence_id": 19,
        "sentence_form": "말 달리면서 부연 모래 바람이 일 때 따라란탄 땀땀~ 꽤 호쾌한  
느낌을 줬다.",
        "opinions": [
            {
                "opinion_id": 1,
                "category": "음향/음악#연출력",
                "opinion_target": "Null",
                "begin": 0,
                "end": 0,
                "polarity": "positive"
            }
        ]
    },
    {
        "sentence_id": 20,
        "sentence_form": "충소리도 실감나고..",
        "opinions": [
            {
                "opinion_id": 1,
                "category": "음향/음악#일반",
                "opinion_target": "충소리",
                "begin": 0,
                "end": 3,
                "polarity": "positive"
            }
        ]
    },
    {
        "sentence_id": 21,
        "sentence_form": "디테일은 확실히 훌륭했다.",
        "opinions": [
            {
                "opinion_id": 1,
                "category": "영화#연출력",
                "opinion_target": "Null",
                "begin": 0,
                "end": 0,
                "polarity": "positive"
            }
        ]
    },
    {
        "sentence_id": 22,
        "sentence_form": "미술적인 면도 나쁘지 않았어.",
        "opinions": [
            {
                "opinion_id": 1,
                "category": "영상/시각 요소#일반",
                "opinion_target": "미술적인 면",
                "begin": 0,
                "end": 6,
            }
        ]
    }
}

```

```

        "polarity": "neutral"
      }
    ]
  },
  {
    "sentence_id": 23,
    "sentence_form": "다만 그 난리를 치고 다니는데 정우성의 셔츠가 너무 깨끗하고 잘
다려져 있는 게 보여서 약간 거슬렸을 뿐. ㅋㅋ",
    "opinions": [
      {
        "opinion_id": 1,
        "category": "영상/시각 요소#연출력",
        "opinion_target": "정우성의 셔츠",
        "begin": 17,
        "end": 24,
        "polarity": "negative"
      }
    ]
  },
  {
    "sentence_id": 24,
    "sentence_form": "지금도 생각난다.",
    "opinions": []
  },
  {
    "sentence_id": 25,
    "sentence_form": "저팔계 같은 `이상한 놈`의 모습.",
    "opinions": []
  },
  {
    "sentence_id": 26,
    "sentence_form": "허영만의 만화에 나오는 모습이 자꾸 오버랩 되네.",
    "opinions": []
  },
  {
    "sentence_id": 27,
    "sentence_form": "아.. 근데 왜 원티드 한 번 더 보고 싶지? -_-",
    "opinions": [
      {
        "opinion_id": 1,
        "category": "OUT OF SCOPE",
        "opinion_target": "Null",
        "begin": 0,
        "end": 0,
        "polarity": "positive"
      }
    ]
  }
]
},
]
}

· text_level
{
  "document_id": "ERRW1903002756.23",
  "domain": "영화",
  "is_test": "N",
  "metadata": {
    "title": "놈놈놈",
    "author": "김마리짱",
    "publisher": "egloos",
    "date": "20080718",

```



```

"url": "http://kimmari02.egloos.com/1876418"
},
"sentence": [
  {
    "sentence_id": 1,
    "sentence_form": "개봉 전에 평이 엇갈려서 보러갈까 말까 망설였는데 요즘 딱히 볼
영화도 없고 그냥 화제작이라는 건 늘 개봉 첫날 구미가 바삭 당기는 법이니까 다녀왔다."
  },
  {
    "sentence_id": 2,
    "sentence_form": "아니나다를까 온갖 대학생, 아줌마, 유모차, 백수, 할머니
할아버지까지 넘쳐나더군. -_-;"
  },
  {
    "sentence_id": 3,
    "sentence_form": "휠체어를 밀고 오신 분도 봤다."
  },
  {
    "sentence_id": 4,
    "sentence_form": "이 정도의 인원이 목요일에 다녀갈 정도면 파급력은 확실할 거
같다."
  },
  {
    "sentence_id": 5,
    "sentence_form": "애초에 기대를 반쯤 접고 봐서 그런지 나쁘지는 않았는데 사실은
좀 흐음.. 했다."
  },
  {
    "sentence_id": 6,
    "sentence_form": "스토리 라인이 너무 빈약하더군."
  },
  {
    "sentence_id": 7,
    "sentence_form": "물론 후반부의 추격씬이랄까 싸우는 광경은 꽤 잘 찍었다."
  },
  {
    "sentence_id": 8,
    "sentence_form": "정말 웨스턴 무비 같다."
  },
  {
    "sentence_id": 9,
    "sentence_form": "이런 장르가 한국영화에는 없었기 때문에 더욱 실감나기도 하고.."
  },
  {
    "sentence_id": 10,
    "sentence_form": "송강호의 '이상한 놈'이 주인공 같은데 평소 캐릭터를 심분 살려서
곳곳에 유머를 배치하려고 애쓴 것도 보이고 물론 나도 몇 번 웃긴 했는데 그냥 피식 정도.."
  },
  {
    "sentence_id": 11,
    "sentence_form": "어떤 건 너무 의도적으로 웃기려고 한 대사라 전혀 우습지가
않았다."
  },
  {
    "sentence_id": 12,
    "sentence_form": "강철중 같은 대사발을 기대하면 필시 실망할 듯.."
  },
  {
    "sentence_id": 13,
    "sentence_form": "그래서 재미없다는 건 아니고 액션이 괜찮았기 때문에 중박
정도라고 생각한다."
  }
]

```

```

    "sentence_id": 14,
    "sentence_form": "이병헌과 정우성의 캐릭터는 그야말로 딱 중간 정도.."
  },
  {
    "sentence_id": 15,
    "sentence_form": "확실히 송강호가 왜 맨 첫번째로 이름이 나왔는지 알 거 같았다."
  },
  {
    "sentence_id": 16,
    "sentence_form": "분명 나이순은 아닌 것 같다. -_-;"
  },
  {
    "sentence_id": 17,
    "sentence_form": "그 중간이라는 게 아쉬운 거지.."
  },
  {
    "sentence_id": 18,
    "sentence_form": "칭찬하고 싶은 건 영화 음악."
  },
  {
    "sentence_id": 19,
    "sentence_form": "말 달리면서 부연 모래 바람이 일 때 따라란탄 땀땀~ 꽤 호쾌한
느낌을 줬다."
  },
  {
    "sentence_id": 20,
    "sentence_form": "충소리도 실감나고.."
  },
  {
    "sentence_id": 21,
    "sentence_form": "디테일은 확실히 훌륭했다."
  },
  {
    "sentence_id": 22,
    "sentence_form": "미술적인 면도 나쁘지 않았어."
  },
  {
    "sentence_id": 23,
    "sentence_form": "다만 그 난리를 치고 다니는데 정우성의 셔츠가 너무 깨끗하고 잘
다려져 있는 게 보여서 약간 거슬렸을 뿐. ㅋㅋ"
  },
  {
    "sentence_id": 24,
    "sentence_form": "지금도 생각난다."
  },
  {
    "sentence_id": 25,
    "sentence_form": "저팔계 같은 `이상한 놈`의 모습."
  },
  {
    "sentence_id": 26,
    "sentence_form": "허영만의 만화에 나오는 모습이 자꾸 오버랩 되네."
  },
  {
    "sentence_id": 27,
    "sentence_form": "아.. 근데 왜 원티드 한 번 더 보고 싶지? -_-;"
  }
],
"opinions": [
  {
    "opinion_polarity": "positive",
    "category": "배우#일반"
  }
]

```

```
{
  "opinion polarity": "negative",
  "category": "시나리오#구성력"
},
{
  "opinion polarity": "negative",
  "category": "시나리오#일반"
},
{
  "opinion polarity": "conflict",
  "category": "영상/시각 요소#연출력"
},
{
  "opinion polarity": "neutral",
  "category": "영상/시각 요소#일반"
},
{
  "opinion polarity": "conflict",
  "category": "영화#연출력"
},
{
  "opinion polarity": "conflict",
  "category": "영화#인지도"
},
{
  "opinion polarity": "conflict",
  "category": "영화#일반"
},
{
  "opinion polarity": "positive",
  "category": "음향/음악#연출력"
},
{
  "opinion polarity": "conflict",
  "category": "음향/음악#일반"
},
{
  "opinion polarity": "conflict",
  "category": "캐릭터#일반"
}
]
```

· 자료 내용 문의: 02-2669-9614