

6가지 감정 분석

게임 인공지능 기말프로젝트

게임소프트학과

C077044 황태훈

PROJECT GOAL

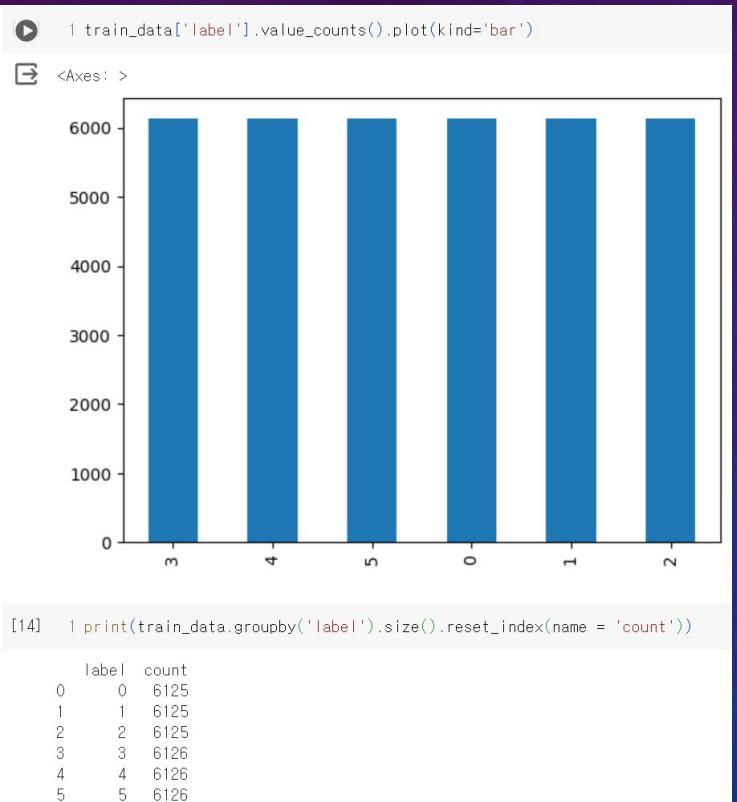
- 감성 분석 학습을 통해 대화문장의 6가지 감정으로 분석
- 대화 문장이 부정적인 문장인지 긍정적인 문장인지 분석

데이터셋 정리

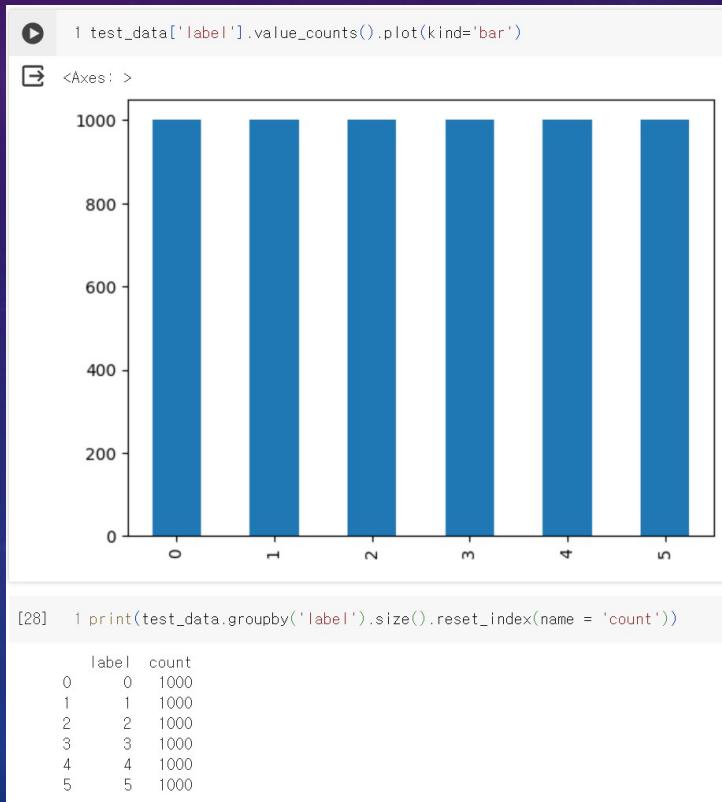
idx	연령	성별	상황키워드	신체질환	성정_대분류	성정_소분류	사람문장1	스팀문장1	사람문장2	스팀문장2	사람문장3	스팀문장3	
1	2	1	청년	여성	진로, 취업, 해당없음	분노	노여워하는일은 왜 하 많이 힘드. 그냥 내가 혼자 해결하기로 했군요 혼자서 해결하기 힘들면 주위에 의논할 사람을 찾아보세요.						
2	3	2	청년	여성	진로, 취업, 해당없음	분노	노여워하는이번 달에 금여가 줄 최대한 지 월급이 줄어든 만큼 소비를 줄일 계획이군요.						
3	4	3	청년	여성	진로, 취업, 해당없음	분노	노여워하는회사에 신 회사 동료 잘 안 맞는스트레스받지 않기 위해선 인간관계에 있어 약간의 거리를 두는 게 좋겠군요.						
4	5	4	청년	여성	진로, 취업, 해당없음	분노	노여워하는직장에서 관련 없는 직장 사람·직장 사람들과 이야기를 해 보겠다고 결심하셨군요.						
5	6	5	청년	여성	진로, 취업, 해당없음	분노	노여워하는얼마 전 입주시하는 상사인 나 항상 먼저 인사하게 되어 화가 나셨군요. 어떻게 하면 신입사원에게 화났음을 표현할 수 있을까요?						
6	7	6	청년	여성	진로, 취업, 해당없음	분노	노여워하는직장에 대 진로에 대; 직장 상사: 업무가 나와 맞지 않아 시간을 베리는 것 같이 느껴지셨군요.						
7	8	7	청년	여성	진로, 취업, 해당없음	분노	노여워하는성인인데 부모님의 부모님께 진로 테스트를 받아보겠다고 부모님께 말씀드릴 생각이군요.						
8	9	8	청년	여성	진로, 취업, 해당없음	기쁨	느긋 퇴사한 지 천천히 라더 좋은 회사는 일자리를 놓치고 싶지 않아 끊임없이 노력하시는군요. 하는 일이 잘 풀리시길 바라요.						
9	10	9	청년	여성	진로, 취업, 해당없음	불안	걱정스러운줄업반이라취업에 대; 응. 느긋한 걱정보다는 느긋한 태도가 낫다고 생각하시는군요.						
10	11	10	청년	여성	진로, 취업, 해당없음	기쁨	느긋 요즘 직장 직장생활이 우리 회사 예비 엄마들에게 정말 꿈의 회사인 것 같네요. 마음이 편하시게어 느긋 취업해야 취업하고 아직 조금 아직은 인생을 즐기고 싶은 마음이시군요.						
11	12	11	청년	여성	진로, 취업, 해당없음	불안	당혹스러운면접에서 그때 어떤 무식 당혹. 그런 면접 상황에 무척 놀라셨군요. 어떻게 답하셨는지 어 당혹스러운면접일이야. 중요한 서. 혹시 과장 혼이 날짜 밖 과장님께 서류의 행방을 물어보는 것이 고민 당혹스러운 얼마나 전회사에서 '코로나바이러스' 면접을 보지 않았으면 더 좋았겠단 생각이 드시는 당황 길을 가다 취업에 대 평소 친한 친지에 관해 물어보셔서 당혹스러운 상황이었군요.						
12	13	12	청년	여성	진로, 취업, 해당없음	불안	당황스러운면접에서 그때 어떤 무식 당혹. 그런 면접 상황에 무척 놀라셨군요. 어떻게 답하셨는지 어 당혹스러운면접일이야. 중요한 서. 혹시 과장 혼이 날짜 밖 과장님께 서류의 행방을 물어보는 것이 고민 당혹스러운 얼마나 전회사에서 '코로나바이러스' 면접을 보지 않았으면 더 좋았겠단 생각이 드시는 당황 길을 가다 취업에 대 평소 친한 친지에 관해 물어보셔서 당혹스러운 상황이었군요.						
13	14	13	청년	여성	진로, 취업, 해당없음	불안	당황스러운면접에서 그때 어떤 무식 당혹. 그런 면접 상황에 무척 놀라셨군요. 어떻게 답하셨는지 어 당혹스러운면접일이야. 중요한 서. 혹시 과장 혼이 날짜 밖 과장님께 서류의 행방을 물어보는 것이 고민 당혹스러운 얼마나 전회사에서 '코로나바이러스' 면접을 보지 않았으면 더 좋았겠단 생각이 드시는 당황 길을 가다 취업에 대 평소 친한 친지에 관해 물어보셔서 당혹스러운 상황이었군요.						
14	15	14	청년	여성	진로, 취업, 해당없음	불안	당황스러운면접에서 그때 어떤 무식 당혹. 그런 면접 상황에 무척 놀라셨군요. 어떻게 답하셨는지 어 당혹스러운면접일이야. 중요한 서. 혹시 과장 혼이 날짜 밖 과장님께 서류의 행방을 물어보는 것이 고민 당황스러운 얼마나 전회사에서 '코로나바이러스' 면접을 보지 않았으면 더 좋았겠단 생각이 드시는 당황 길을 가다 취업에 대 평소 친한 친지에 관해 물어보셔서 당혹스러운 상황이었군요.						
15	16	15	청년	여성	진로, 취업, 해당없음	당황	당황 어제 합격 잘못 발송·물론 사과·대기업의 실수 때문에 실망감을 느끼시는군요.						
16	17	16	청년	여성	진로, 취업, 해당없음	당황	당황 나 오늘 무슨 일 아버스 타고 카드가 없어서 너무 당황스러웠겠어요.						
17	18	17	청년	여성	진로, 취업, 해당없음	당황	당황 이번에 직 이직하신 정말 다시 새 직장에서 만나고 싶지 않은 사람들을 만나서 너무 스트레스로 고생하는군요.						
18	19	18	청년	여성	진로, 취업, 해당없음	당황	마비된 마비된 오늘 회사 무슨 실수·중요한 발·준비 많이 하셨는데 험드셨겠어요.						
19	20	19	청년	여성	진로, 취업, 해당없음	슬픔	마비된 마비된 요즘 취업 온몸에 마·현재 취업 갈수록 힘들어지는 취업에 대해 생각하실 때마다 온몸이 아파요. 아·근이 많아·슬퍼하·말·안 해도 될 아근을 하는 것 같아 더 스트레스받으시는 것						
20	21	20	청년	여성	진로, 취업, 해당없음	슬픔	마비된 마비된 아침에 직 이직하신 정말 다시 새 직장에서 만나고 싶지 않은 사람들을 만나서 너무 스트레스로 고생하는군요.						
21	22	21	청년	여성	진로, 취업, 해당없음	슬픔	마비된 마비된 오늘 회사 무슨 실수·중요한 발·준비 많이 하셨는데 험드셨겠어요.						
22	23	22	청년	여성	진로, 취업, 해당없음	슬픔	마비된 마비된 아침에 직 이직하신 정말 다시 새 직장에서 만나고 싶지 않은 사람들을 만나서 너무 스트레스로 고생하는군요.						
23	24	23	청년	여성	진로, 취업, 해당없음	기쁨	마비된 마비된 아침에 직 이직하신 정말 다시 새 직장에서 만나고 싶지 않은 사람들을 만나서 너무 스트레스로 고생하는군요.						
24	25	24	청년	여성	진로, 취업, 해당없음	기쁨	만족스러운우리 회사·사내 분위·즐거운 일·원하는 진로에 맞춰서 취업해서 더욱 만족스럽군요.						
25	26	25	청년	여성	진로, 취업, 해당없음	기쁨	만족스러운우리 회사에서 시험 결과·열심히 준 준비한 만큼 점수가 잘 나와서 만족스러우시겠어요.						
26	27	26	청년	여성	진로, 취업, 해당없음	상처	만족스러운우리 회사에서 시험 결과·열심히 준 준비한 만큼 점수가 잘 나와서 만족스러우시겠어요.						
27	28	27	청년	여성	진로, 취업, 해당없음	상처	배신당한 배신당한 면접관에게 연봉과 실 회사 내규·신입 연봉체계가 다르시군요.						
28	29	28	청년	여성	진로, 취업, 해당없음	상처	배신당한 배신당한 면접관에게 연봉과 실 회사 내규·신입 연봉체계가 다르시군요.						
29	30	29	청년	여성	진로, 취업, 해당없음	상처	배신당한 배신당한 면접관에게 연봉과 실 회사 내규·신입 연봉체계가 다르시군요.						
30	31	30	청년	여성	진로, 취업, 해당없음	상처	배신당한 배신당한 면접관에게 연봉과 실 회사 내규·신입 연봉체계가 다르시군요.						
31	32	31	청년	여성	진로, 취업, 해당없음	당황	배려진 배려진 계속 취업·내가 능력·취업 실태로 인해 자신감이 많이 떨어진 것 같네요. 힘내시고 부끄러운 부끄러운 이번에 중·중소기업·요즘같이 '오래 준비하고 공부해서 취업하셨군요.'						
32	33	32	청년	여성	진로, 취업, 해당없음	분노	분노 저번 주에 친구분과 친구가 위 친구의 위로에도 마음이 다스려지지 않는군요.						
33							요즘 청년 청년 실업·기업들이 기업 채용을 늘려주는 정책이 생겼으면 좋겠다고 생각하고 있다.						

document label
 퇴사한 지 얼마 안 됐지만 천천히 직장을 구해보려고. 0
 요즘 직장생활이 너무 편하고 좋은 것 같아! 0
 취업해야 할 나이인데 취업하고 싶지가 않아. 0
 우리 회사는 정말 사내 분위기가 좋아. 0
 오늘 내가 다니는 회사가 참 좋은 직장이라는 생각이 들었어. 0
 회사에서 전공시험을 봤어. 오늘 시험 결과가 나왔어. 0
 오늘 입사 면접을 봤어. 0
 회사에서 나를 참 신뢰하는 것 같아. 그건 기분이 좋아. 0
 먼저 취업한 선배가 면접 비결을 알려줬어. 곧 면접 보러 가는데 든든해. 0
 직장 상사로부터 칭찬을 받았는데 너무 신이 나! 0
 취업해서 너무 신이 나! 0
 오늘 사장님께 칭찬받아서 매우 기뻐. 0
 지난주에 정말 가고 싶던 회사에 원서 접수했는데 합격했어. 매우 기뻐. 0
 나 드디어 원하는 회사에 취업했어! 0
 나 오늘 네이버 인턴 면접 봤는데 내가 봐도 잘 본 것 같아. 0
 결혼한 지 삼 년 만에 아이를 낳았어. 진짜 기뻐. 0
 그녀가 나에게 사귀자고 고백했어. 나도 원하고 있었는데 지금 기분이 너무 좋아. 0

데이터셋 분석



훈련 데이터셋



테스트 데이터셋

데이터셋 패딩 전

불용어



단어 집합(vocabulary)의 크기 : 9711

등장 빈도가 2번 이하인 희귀 단어의 수:

단어 집합에서 희귀 단어의 비율: 45.1652

전체 등장 빈도에서 희귀 단어 등장 빈도

[26]

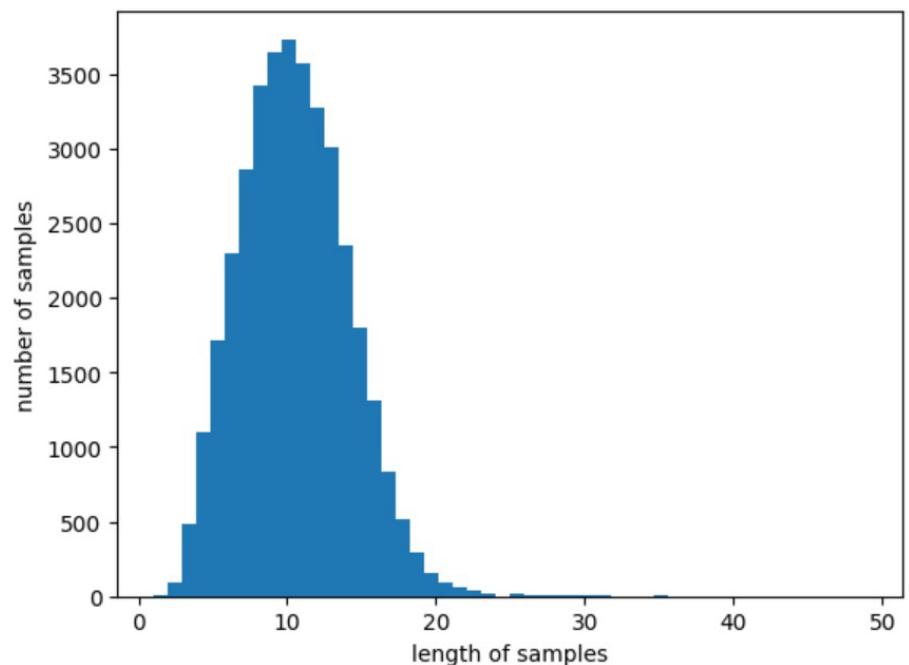
```
1 vocab_size = total_cnt - rare_cnt + 1  
2 print('단어 집합의 크기 : ',vocab_size)
```

단어 집합의 크기 : 5326

데이터셋 패딩

```
1 print('문장의 최대 길이 :',max(len(review) for review in X_train))
2 print('문장의 평균 길이 :',sum(map(len, X_train))/len(X_train))
3 plt.hist([len(review) for review in X_train], bins=50)
4 plt.xlabel('length of samples')
5 plt.ylabel('number of samples')
6 plt.show()
```

문장의 최대 길이 : 49
문장의 평균 길이 : 10.432971458112263



1 max_len = 17

2 below_threshold_len(max_len, X_train)

전체 샘플 중 길이가 17 이하인 샘플의 비율: 96.57715016461242

17

96%

LSMT RESULT

실패

정확도
16.67%

테스트 정확도 : 0.1667

```
▶ 1 sentiment_predict('기쁨')
  2 sentiment_predict('놀람')
  3 sentiment_predict('분노')
  4 sentiment_predict('불안')
  5 sentiment_predict('혐오')
  6 sentiment_predict('슬픔')
```

```
➡ 1/1 [=====] - 0s 50ms/step
  1.0
  1/1 [=====] - 0s 37ms/step
  1.0
  1/1 [=====] - 0s 37ms/step
  1.0
  1/1 [=====] - 0s 39ms/step
  1.0
  1/1 [=====] - 0s 37ms/step
  1.0
  1/1 [=====] - 0s 36ms/step
  1.0
```

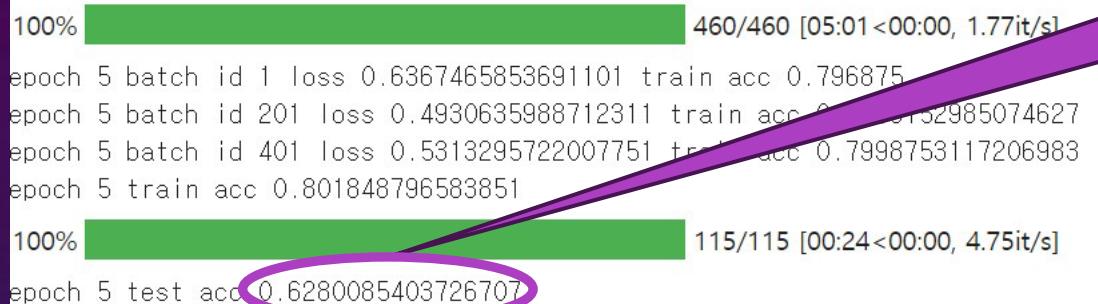
```
[46] 1 sentiment_predict('나 오늘 합격했어')
      1/1 [=====] - 0s 22ms/step
      1.0

[47] 1 sentiment_predict('짜증나! 오늘 과제 너무 많아')
      1/1 [=====] - 0s 22ms/step
      1.0

[48] 1 sentiment_predict('앗, 깜짝이야!')
      1/1 [=====] - 0s 38ms/step
      1.0

[49] 1 sentiment_predict('오늘 큰이모께서 돌아가셨어!')
      1/1 [=====] - 0s 46ms/step
      1.0
```

KOBERT RESULT



정확도: 62.8%

```
1 predict('기쁨')  
2 predict('놀람')  
3 predict('분노')  
4 predict('불안')  
5 predict('혐오')  
6 predict('슬픔')
```

▶>> 입력하신 내용에서 기쁨이 느껴집니다.
100.00% 확률로 긍정 문장입니다.
>> 입력하신 내용에서 놀람이 느껴집니다.
100.00% 확률로 부정 문장입니다.
>> 입력하신 내용에서 분노가 느껴집니다.
100.00% 확률로 부정 문장입니다.
>> 입력하신 내용에서 불안이 느껴집니다.
100.00% 확률로 부정 문장입니다.
>> 입력하신 내용에서 놀람이 느껴집니다.
100.00% 확률로 부정 문장입니다.
>> 입력하신 내용에서 슬픔이 느껴집니다.
100.00% 확률로 부정 문장입니다.

- [40] 1 predict('나 오늘 합격했어 .')
>> 입력하신 내용에서 기쁨이 느껴집니다.
100.00% 확률로 긍정 문장입니다.
- [41] 1 predict('짜증나! 오늘 과제 너무 많아 ')
>> 입력하신 내용에서 분노가 느껴집니다.
100.00% 확률로 부정 문장입니다.
- [42] 1 predict('앗, 깜짝이야 !')
>> 입력하신 내용에서 기쁨이 느껴집니다.
77.03% 확률로 긍정 문장입니다.
- ▶ 1 predict('오늘 큰이 모께서 돌아가셨어 !')
>> 입력하신 내용에서 슬픔이 느껴집니다.
100.00% 확률로 부정 문장입니다.

정확도 약 62% 감정 분석은
성공했지만
문장의 긍정적인 문장인지
부정적인 문장인지의
분석은 결과가 애매합니다.

NEW DATASET KOBERT RESULT

새로운 데이터셋

	label	count
0	0	1386
1	1	1386
2	2	1386
3	3	1386
4	4	1386
5	5	1386

기존 데이터셋

	label	count
0	0	6126
1	1	6126
2	2	6126
3	3	6126
4	4	6126
5	5	6126

4.4배



정확도: 95.4%

```
1 predict('기쁨')
2 predict('놀람')
3 predict('분노')
4 predict('불안')
5 predict('혐오')
6 predict('슬픔')
```

>> 입력하신 내용에서 **혐오**가 느껴집니다.
100.00% 확률로 부정 문장입니다.

>> 입력하신 내용에서 **놀람**이 느껴집니다.
100.00% 확률로 부정 문장입니다.

>> 입력하신 내용에서 **슬픔**이 느껴집니다.
100.00% 확률로 부정 문장입니다.

>> 입력하신 내용에서 **불안**이 느껴집니다.
100.00% 확률로 부정 문장입니다.

>> 입력하신 내용에서 **분노**가 느껴집니다.
100.00% 확률로 부정 문장입니다.

>> 입력하신 내용에서 **기쁨**이 느껴집니다.
100.00% 확률로 긍정 문장입니다.

```
1 predict('나 오늘 합격했어 .')
```

>> 입력하신 내용에서 **기쁨**이 느껴집니다.
100.00% 확률로 긍정 문장입니다.

```
1 predict('짜증나! 오늘 과제 너무 많아')
```

>> 입력하신 내용에서 **분노**가 느껴집니다.
100.00% 확률로 부정 문장입니다.

```
1 predict('앗, 깜짝이야!')
```

>> 입력하신 내용에서 **혐오**가 느껴집니다.
100.00% 확률로 부정 문장입니다.

```
1 predict('오늘 큰이 모께서 돌아가셨어 !')
```

>> 입력하신 내용에서 **놀람**이 느껴집니다.
100.00% 확률로 부정 문장입니다.

DISCUSSION

- What I learned from this project

인공지능 학습에 따라서 감성분석에 성공여부가 차이가 난다는 것을 배우게 되었습니다.

학습 프로그램 뿐만 아니라 데이터양도 무시 못한다는 사실도 알게 되었습니다.

- How to improve the results

데이터셋을 더 많이 모았으면 정확도가 개선될 것 같습니다.