

데이터분석캡스톤디자인 주차 결과 발표

CONTENTS. |

01

감성 분류 모델

02

Word2vec 설계

03

감성사전 구축결과

04

다음주 수행계획

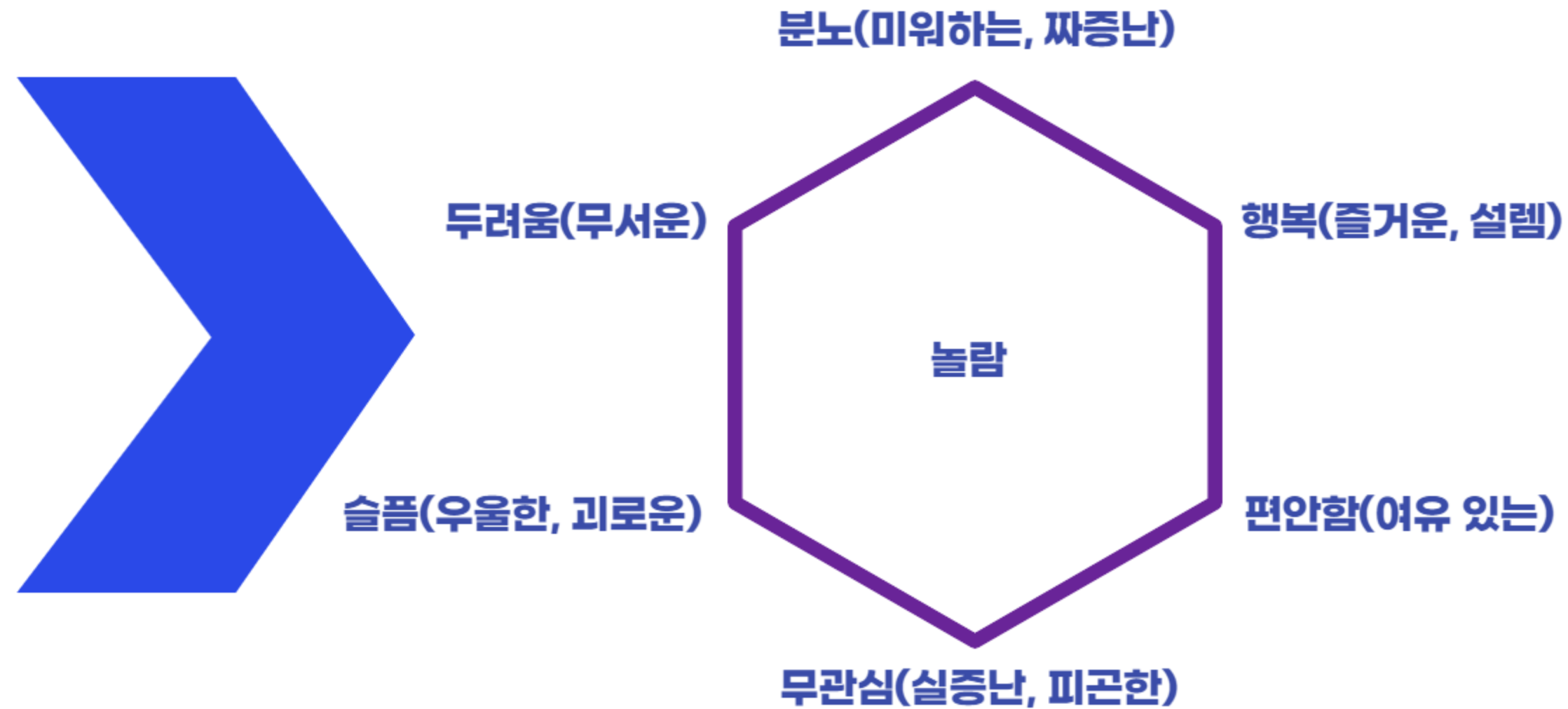
감성 분류 모델

01. Russell(1980) 모델(4가지)

- 1)고각성 긍정정서: 신이 난, 흥분한, 놀라운,
아주 기뻐하는, 행복한
- 2)저각성 긍정정서: 기쁜, 만족스러운, 흡족한,
고요한, 차분한, 여유 있는
- 3)고각성 부정정서: 불안해 하는, 두려워 하는,
긴장한, 화난, 괴로워 하는, 짜증 난
- 4)저각성 부정정서: 비참한, 우울한, 슬픈, 따
분한, 의기소침한

02. Tellegen-Watson Clark 모델(8가지)

- 1)강한 활성화: 놀란
- 2)강한 긍정적인 영향: 흥분된, 즐거운
- 3)유쾌한: 행복, 유쾌한
- 4)약한 긍정적인 영향: 편안한, 평온한
- 5)약한 활성화: 무관심한, 노곤한
- 6)약한 부정적인 영향: 권태로운, 피곤한
- 7)강한 부정적인 영향: 두려운, 초조한
- 8)불쾌한: 실증난, 슬픔



Word2Vec 모델 설계

01. CBOW vs. Skip-gram 문맥(주변단어)을 통해 감정(중심단어)를 파악하는 것이 목적이기 때문에 CBOW 활용

02. size 100차원의 벡터로 설정 **03. window** = 5 **04. iter** 10회 반복 **05. min_count** 최소 빈도 수는 3

```
[13] from gensim.models.word2vec import Word2Vec
```

```
[14] model = Word2Vec(tagging_kkma,  
                     size = 100,  
                     sg = 0,  
                     window = 5,  
                     iter = 10,  
                     min_count = 3)
```

감성사전 구축 결과

01. 감성별 키워드 결정

감정별 주요 키워드를 입력해보면서 포함된 감정을 아우를 수 있는 최종 키워드 결정

1) 분노(미워하는, 짜증난)

미움, 미워하(다)

2) 두려움(무서운)

두렵(다), 무섭(다)

3) 슬픔(우울한, 괴로운)

슬픔, 슬프(다), 우울

4) 무관심(실증난, 피곤한)

무관심, 질리(다), 피곤

5) 편안함(여유 있는)

편안, 여유

6) 행복한(즐거운, 설레는)

행복, 즐기(다), 설레(다)

7) 놀람

놀라(다)

감성사전 구축 결과

02. 감정별 감성사전 구축

```
[18] df = pd.DataFrame(model.wv.most_similar("질리", topn = 30), columns = ['단어', '유사도'])
```

```
print(df)
```

	단어	유사도
0	흐트러지	0.593322
1	센치	0.588427
2	말랑	0.575145
3	둥글	0.569930
4	불안해하	0.567073
5	구차	0.565793
6	조급	0.557273
7	초조	0.554945
8	심하	0.547586
9	헛갈리	0.545770
10	심각	0.528727
11	불쌍	0.522246
12	이쁘	0.521392
13	귀찮	0.520387
14	피곤	0.518527
15	지독	0.517094
16	공평	0.512941
17	불편	0.509184
18	흐릿	0.509165
19	굴하	0.508533

가장 유사도가 높은 순서로 출력

유사도 0.5 이상만 반영

두가지 이상의 감정에 중복해서 포함되는
단어는 배제 예정

감성사전 구축 결과

02. 감정별 감성사전 구축

단어	유사도	단어	유사도	단어	유사도	단어	유사도	단어	유사도
0	무섭 0.620404	0	두렵 0.620404	0	소중 0.685871	0	설레이 0.855638	0	아픔 0.644182
1	겁 0.576229	1	버겁 0.594166	1	감사 0.610717	1	떨리 0.716911	1	흔적 0.599066
2	변하 0.565953	2	어렵 0.589658	2	사랑 0.599444	2	두근거리 0.628561	2	추억 0.555937
3	바뀌 0.564486	3	두려워지 0.585994	3	부족 0.576715	3	들뜨 0.578696	3	그리움 0.547894
4	버겁 0.562713	4	이르 0.545584	4	충분 0.551058	4	울렁이 0.508563	4	바람 0.514567
5	잊히 0.556402	5	웃기 0.543546	5	고맙 0.518791	5	두근대 0.504997	5	외로움 0.509749
6	슬퍼지 0.542556	6	실 0.537732	6	간직 0.514887	6	순수 0.495807	6	희망 0.507887
7	멀 0.514455	7	늦 0.535819	7	완벽 0.505423	7	허전 0.490726	7	기쁨 0.505990
8	언제쯤 0.510707	8	아깝 0.535774	8	선명 0.498037	8	두근두근 0.487438	8	미움 0.499673
9	다가서 0.500529	9	지겹 0.534636	9	미안 0.496420	9	간절 0.481502	9	어둠 0.493628
10	달라지 0.494929	10	그리워지 0.531306	10	힘들 0.484871	10	편안 0.471033	10	시련 0.489305
11	도망가 0.490702	11	우습 0.529250	11	이해 0.473115	11	써 0.468229	11	나날 0.484386
12	혹시나 0.487931	12	즐겁 0.528765	12	불행 0.471442	12	느끼 0.461515	12	짐 0.478619
13	느리 0.486784	13	부럽 0.522126	13	함께 0.470749	13	느껴지 0.456078	13	괴물 0.474811
14	궁금하 0.481020	14	길 0.520334	14	초라 0.455377	14	상쾌 0.454697	14	계절 0.468772
15	서두르 0.473798	15	재밌 0.509628	15	아파하 0.451121	15	어지럽 0.441861	15	고통 0.464594
16	미워지 0.465566	16	바뀌 0.509203	16	따뜻 0.446898	16	끌리 0.433886	16	터널 0.463167
17	그리워지 0.463615	17	변하 0.500202	17	아름답 0.445092	17	신기 0.427411	17	오감 0.460739
18	힘들 0.460142	18	반갑 0.500040	18	영원 0.444509	18	피어나 0.418970	18	침묵 0.445471
19	나아지 0.459252	19	느리 0.494922	19	기억 0.438712	19	뒤 0.412475	19	깨끗이 0.445194

0	허전	0.734751
1	투명	0.723228
2	흐릿	0.709014
3	불안	0.685982
4	조급	0.678577
5	초라	0.678541
6	아련	0.674776
7	아득	0.673303
8	먹먹	0.672094
9	쓸쓸	0.667074
10	무력	0.661226
11	무색	0.660500
12	선명	0.658637
13	간절	0.657008
14	서운	0.655005
15	나른	0.641983
16	무심	0.640441
17	여전	0.639391
18	말랑	0.637502
19	곳허	0.634005

다음주 수행계획

1. 구축한 감성사전을 활용하여 곡별 감정 라벨링

2. 딥러닝 분류 모델 결정 및 학습 시작

