

업무 보고서

2021.06.01

인공지능

오종민

목차

1. 업무 개요

2. 감정 분류 방법론

3. 한국어 감정 분류

4. 영어 감정 분류

5. 추가 방안

6. 기타 업무

7. 업무 관리

가사 & 한국어의 한계

가사

가사데이터에는 비유적 표현이 많고 직설적이지 않고 돌려 말하는 경우가 대부분

사랑하는 사람을 이젠 볼 수 없어요. -> 헤어졌다.

좋아했던 사람과 만날 수 없네요 -> 슬프다.

가슴에 멍이 든 것 같아요 -> 아프다.

너가 없으면 죽을 것만 같아 -> 사랑한다.

너를 잃기 싫어 -> 사랑한다.

부은 눈으로 잠에 들었다. -> 울었다.

가사 전체를 보았을 때 느껴지는 감정이 여러가지다.

-> 정확도를 정량적으로 측정하기가 어렵다.

-> 어느 정도 납득이 되는 결과인지 아닌지를 판단 해야함.

한국어 자연어 처리

1. 접사 추가에 따른 의미 발생
 - ‘사과’ + 를 은 목적어가 되고
 - ‘사과’ + 가 는 주어가 된다.교착어인 한국어의 대표적인 특징
2. 유연한 단어 순서
 - 나는 밥을 먹으러 간다.
 - 밥을 먹으러 나는 간다.
 - 나는 간다, 밥을 먹으러
3. 모호한 띄어쓰기 규칙
4. 평서문과 의문문의 차이 없음, 주어 부재

가사 감정 분류

필요성 :

리뉴얼 방안 따라 음원마다 가사에서 느껴지는 감정들을 추천시스템의 Feature(특징)로 사용할 수 있다.

- 사용자의 현재 감정에 맞는 채널 추천
- 감정이 비슷한 음원들을 묶어 채널 생성

개요 :

가사 텍스트에서 느껴지는 감정을 분류하는 인공지능 개발

10개의 감정 클래스로 구분 :

사랑해 , 불안해요, 신나요 , 편안해요 , 졸려요 , 설레요 , 그리워요 , 슬퍼요 , 싱숭생숭, 화나요

데이터 셋 :

3311	4179	주현미-그때 그밤은 따뜻했었네	[7]	204482.0	한때는 나의 곁에 앉아 있을 때WrWn행복을 느낀다고 말하던 당신 WrWn지금은 그 ...
3312	4180	주현미-꽃 한송이	[8]	135283.0	그 누가 껴있나 한 송이 외로운 꽃WrWn시들은 꽃 송이가 향흔빛에 애담구나WrWn...
3313	4183	주현미-올고넘는 박달재	[8, 7]	129369.0	천등산 박달재를 올고 넘는 우리 임아WrWn물항라저고리가 굶은비 에 젖는구려WrWn왕...
3314	4184	주현미-정말 좋았네	[7, 8]	40162.0	1.WrWn 사~랑 그 사랑이 정말 좋았네WrWn 세~월 그 세월이 가는 줄도 모르...
3315	4185	주현미-짝사랑	[6]	28660.0	마주치는 눈빛이 무엇을 말하는지 WrWn난아직 몰라 난정말 몰라 WrWn가슴만 두근...

3316개의 가사데이터와 10가지로 된 라벨링

기존 데이터 셋의 문제점

여러 개의 라벨데이터

- 한 노래 가사당 감정이 1~3개 까지 달려있다.
- 분류 학습하기 매우 복잡한 구조 *그러나 가사에 감정이 하나만 존재하는 것도 이상함.

불균형 문제

- 그리워요, 슬퍼요, 싱송생송의 비율이 너무 높음

분류 라벨이 너무 많이 존재 : 10개의 감정

- 분류 개수가 많아 질 수록 정확도는 떨어질 수 밖에 없다.
- 그렇다고 해서 10가지의 감정이 뚜렷하게 나타나지도 않는다.

문서(가사) 단위가 너무 길다.

- 평균 702자, 최대 3903자
- * 많은 단어들을 고려하여 감정을 결정하기엔 어려움이 있다.
- * kaggle data 중에 spotify 에서 배포한 가사에 대한 긍/부정 정도 데이터 셋으로.
 긍/부정만 판단하는 모델에도 BERT로 학습시켰을 때 63% 정확도를 보였음.

학습 데이터양이 너무 적다

- 3316 개의 가사 데이터

결론 : 이 데이터 셋을 사용하기엔 불가능하다고 판단함.

데이터 셋을 새로 구축하기로 결정. (지훈님과 팀장님의 걱정...)

데이터 셋 구축

데이터 셋을 새로 구축하고 결정.

방향을 가사 **전체** 분류 모델에서 가사 **문장** 분류 모델로 목적을 변경함.

필요 이유 : 가사 도메인에 대한 데이터 셋이 필요함.

(한국어 단발성 대화 데이터 셋을 사용해 보았지만 가사 감정에 사용하기엔 방향이 다름)

가사 단위는 너무 길어서 문장 단위 데이터 라벨링 필요.

문장 당 하나의 감정이 라벨링 된 데이터 필요.

가사 전체 단위에서 문장단위로 변경

- 데이터에 포함 되어있는 개행문자('w'n,w'r'n')를 기준으로 문장단위로 분할
- 3316개의 데이터가 -> 66000개의 문장데이터로 분할됨

분할된 문장 데이터 전처리

- 특수 문자 제거
- 의미 전달이 부족한 짧은 문장 제거 (글자 수 10개 이하 제거)
- 문장 앞뒤 공백제거
- 중복 문장 제거(후렴구 반복 등장 방지) * 띄어쓰기를 전부 없애고 중복을 제거함.

한글/영어 구분

- 오직 영어만 있는 경우는 따로 분리함.
- 한국어가 하나라도 포함 되어있는 문장만 남김
- 66000 -> 55000 개

언니 동생으로 부르는데 맞는 일인가요..??
그냥 내 느낌일뿐겠지?
아직너무초기라서 그런거죠?
유치원버스 사고 났다던데
근데 원래이런거맞나요
남자친구가 떠날까봐요
이거 했는데 허리가 아플수도 있나요? ;
내가불안해서꾸는걸까..
일주일도 안 남았당...πππ
약은 최대한 안먹으려고 하는데좋은 음식있나요?0
구직활동 하면서 남은시간은 뭘로 활용해야 되지..

-> 공포

한국어 단발성 대화 데이터 셋

데이터 셋 구축

* 기존 라벨로 어느 정도 라벨링을 진행하면서
감정 종류를 재정의했습니다.

6가지 감정으로 축약

- 빈도 수가 너무 적게 나오는 감정들 제거 (졸리다 , 싱숭생숭)
- 비슷한 감정은 묶음으로 처리 (두려움 , 그리움 -> 공허)

전혀 감정이 없는 문장은 제외

- 0 으로 라벨링

- 떨림 , 사랑 , 즐거움 , 슬픔 , 분노 , 공허

감정 분류 라벨링 규칙을 정의하여

포괄적으로 분류하고 라벨의 특징을 포함

문장의 맥락을 고려하여 분류 태깅함.

1. 떨림

- 생각한다, 기분, 분위기, 소원, 어색하다
- 처음 봤을 때 (시각적인 표현), 시선표현, 몰래
- 좋아해 주길 바라는, 우리 같이, ~하고 싶어, ~ 할 꺼야
- 느낌 표현, 같이 걸었다, 기다린다, 달콤하다 부드럽다, 새롭다
- (잠)에 대한 표현, 눈빛, 만나자, 옷을 입는다, 미소, 선물

2. 사랑

- = LOVE, 사랑, 좋아해, 반했다. 아끼다, 고백
- 고마움, 보고싶어, 아름다워, 심장, 예쁘다, 안아줘
- 옆에 있어줘, 다가와, 너만 있으면 돼, 널 위해 ~해줄게, 소중한다, 지켜줄게

3. 즐거움

- 웃는다, 즐겁다, 신난다, 뛰자, 달리자, 춤추자, 편하다, 하늘
- 행복하다, 손들어(hands up), 소리쳐, 박수쳐, 축하, 랩, 힙합
- 술먹자(좋아서 먹는 술) 무대의 관한 내용, 게임, 쉬자, 휴식
- 희망, 꿈(Dream)만 같다, 돈 벌자, 시끄럽다, 장난, 부자, 랜잡아

4. 슬픔

- 아프다, 힘들다, 후회된다, 미련, 미워하다, 미안하다, 애통하다, 취했다(술퍼서), 가엾다
- 울었다, 이별, 헤어졌다, 우울하다, 잊다, 한숨을 내쉬다, 아쉽다, 모른다, 걱정, 무겁다
- 비, 한심하다,

5. 분노

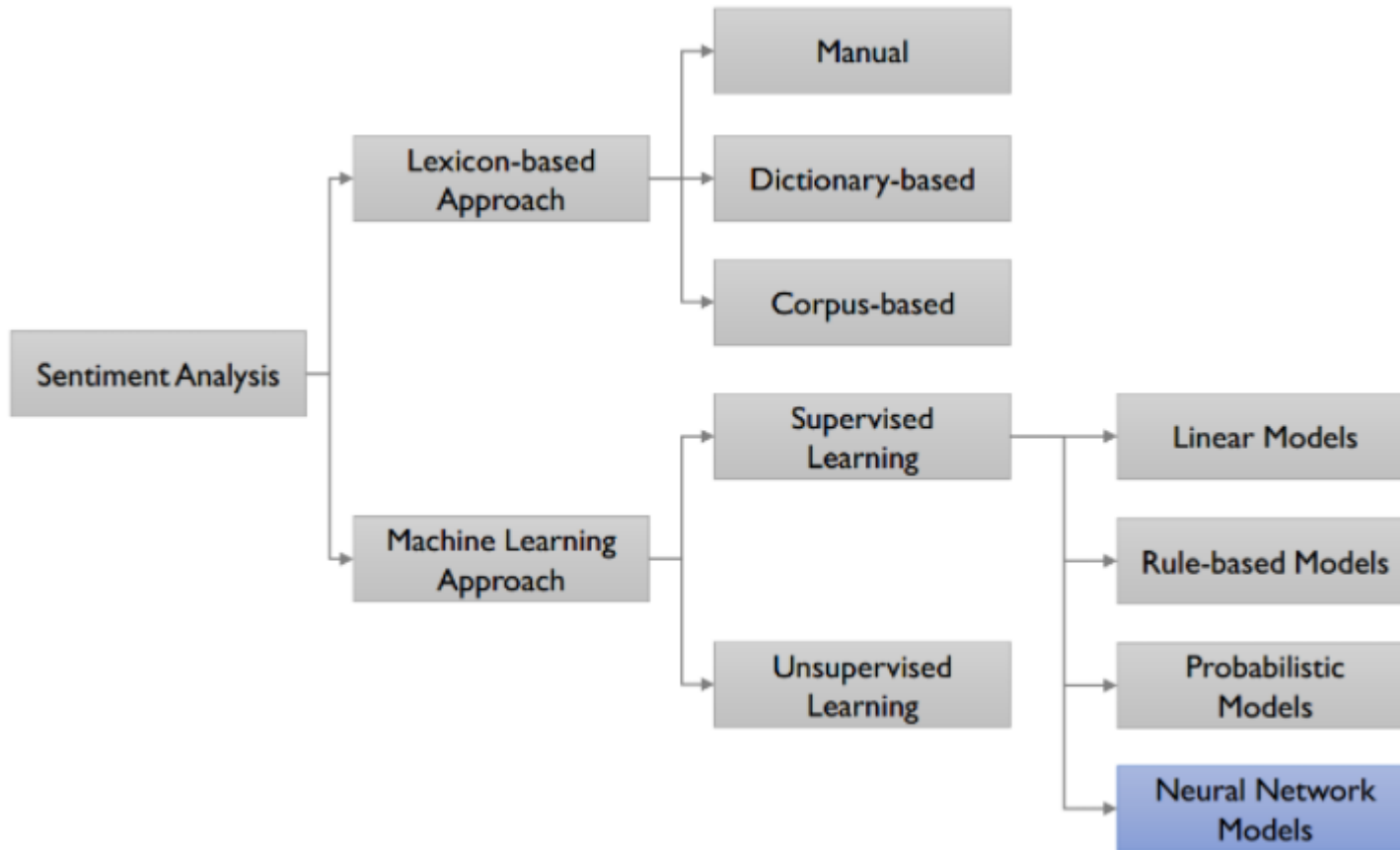
- 패배, 악몽, 싫어, 잔인하다, 거짓말, 비속어, 귀찮다, 복잡하다, 그만하다
- 내버려둬, 실수, 화내다, 찢다, 지겹다, 어렵다, 불쾌하다,
- ~하지마, 지루하다, 피곤하다, 욕먹다, 소름끼친다, 가라(떨어져라 나한테서)

6. 공허

- 외롭다, 멀어지다, 그림자, 빈자리, 두렵다, 무섭다, 버리다, 상실감, 무심한
- 할 수 있는 게 없다, 해줄 게 없다, 모자라다, 부족하다. 되돌리다, 어둠, 아무것도
- 도망쳐, 돌아와줘, 쓸쓸하다, 차갑다, 무심하다, 혼자, 기억, 떠난다, 추억
- 지나가다, 허탈하다, 정리하다, 구해줘, 배고프다

라벨링 규칙 정의

감정 분류 방법론



가사 감정 분류 방법론

단어 사전 기반 가사 문장 감정 분류

The image displays four panels, each representing a different emotion category: thrill, love, joy, and sad. Each panel contains a grid of related Korean words and a '+ 데이터 추가' (Add Data) button.

- thrill (41):** 생각하다, 선물, 웃, 분위기, 소원, 달콤, 생각, 보다, 부드럽다, 맘, 모습, 마음, 잠, 눈빛, 향기, 함께, 같이, 멀리다, 미소, 우연히, 아침, 처음, 늘, 밤, 잠, 위하다, 기다리다, 만나다, 걸다, 느끼다, 달다, 바라보다, 몰래, 빛나다, 다가오다, 설레다.
- love (30):** 사랑, 사랑은, 사랑해, 심장, 마음, 맘, 입술, 고백, 지키다, 좋아하다, 미지다, 반하다, 아끼다, 아름답다, 좋다, 안다, 예쁘다, 가득, 벽차, 벽차다, 영원히, 소중, 언제나, 걸, 내 사랑, love, mellow, together, 눈부시다, 뜨겁다.
- joy (51):** 하늘, 무대, 비트, 음악, 웃음, 웃다, 부르다, 놀다, 즐기다, 추다, 흔들다, 달리다, 말기다, 움직이다, 향하다, 날아가다, 꾸다, 뜨겁다, 편하다, 꿈같다, 재미있다, 멋있다, 재있다, dance, 후회 없다, party, good, 신나다, 희망, 별다, 시끄럽다, 장난치다, 장난.
- sad (28):** 헤어지다, 흐르다, 흘리다, 아프다, 슬프다, 밉다, 힘겹다, 아쉽다, 가엾다, 미련, 서글프다, 애통, 어지럽다, 한숨, sorry, 한심, 한없이, 어차피, 눈물, 이별, 후회, 불안, 슬픔, 모르다, 지치다, 잊다, 울다, 지우다.

IBM Watson Developer Cloud의 API 를 사용하여 감정 단어를 어느정도 정리하고 테스트를 해보았습니다.

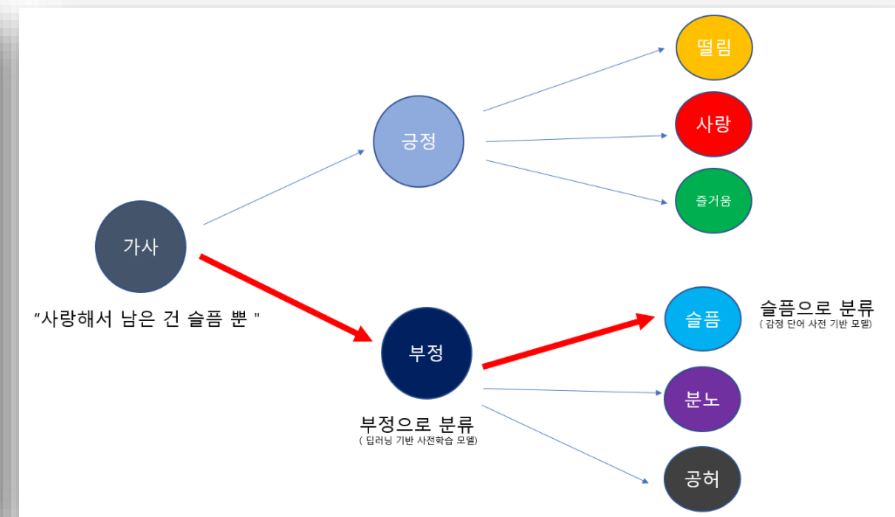
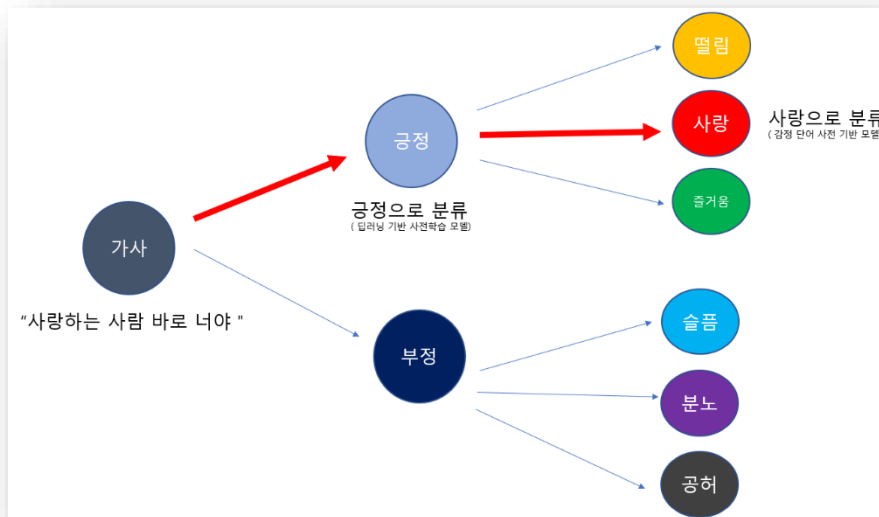
여기서 테스트 해본 결과가 생각보다 잘 나와서 감정 단어 사전 기반 모델을 구축하기로함.

그런데

단어들의 우선순위를 판별하기 어려움, '사랑해'와 '슬프다'가 같이 나왔을 때 뭐가 더 큰 감정인지 판단 할 수 없음.
'널 사랑하지 않아'를 사랑으로 분류하기에 부정어 처리를 하지 못하는 문제 발생

가사 감정 분류 방법론

딥러닝 기반 + 사전 기반 = 하이브리드 기법



6가지 클래스 분류문제를 바로 하는 것 보다.

중간에 긍/부정 필터를 두어서 **문장의 전체적인 문맥**을 파악하고
필터를 거친 문장들은 각각의 **단어들로 감정 정도를 판단하여 점수를 합산하는 모델**

머신 기반 감정 분류

Machine-based Approach

장점 : 머신 기반이기 때문에 **성능이 좋은 모델들을 사용할 수 있다.**

단점 : 특징을 찾아내서 분류하기 때문에 비슷한 세부 감정을 분류할 때 정확도가 떨어 질 수 있다.

머신러닝 기술이 발달하면서 어휘 기반의 감성 분석 방법보다는 머신러닝 모델 기반의 감성 분석이 많이 수행되고 있습니다. 특히 지도 학습(Supervised Learning) 기반의 방법이 많이 시도되고 있습니다.

RNN(순환 신경망 모델)

- 은닉층의 노드에서 활성화 함수를 통해 나온 결과값을 출력층 방향으로도 보내면서, 다시 은닉층 노드의 다음 계산의 입력으로 보내는 특징
- 짧은 시퀀스에 대해서만 효과를 보이는 단점.

LSTM(장단기 메모리)

- LSTM은 은닉 상태(hidden state)를 계산하는 식이 전통적인 RNN보다 조금 더 복잡해졌으며 셀 상태(cell state)라는 값을 추가하였습니다 (중요한 값만 남기고 나머지는 버린다.)

GRU(게이트 순환 유닛)

- LSTM의 장기 의존성 문제에 대한 해결책을 유지하면서, 은닉 상태를 업데이트하는 계산을 줄였습니다.
- LSTM 단순화

BiLSTM

- LSTM 에 어텐션 기법을 추가한 방법
- 예측해야할 단어와 연관이 있는 입력 단어 부분을 좀 더 집중(attention)해서 봄

BERT(양방향 인코더)

- 사전 훈련된 워드 임베딩
- 사전 훈련된 언어 모델
- 마스크 언어 모델

머신 기반 감정 분류(BERT)

Sentence binary classification

문맥을 고려하는 신경망 모델을 사용해야함.

최근 성능이 좋은 사전학습 모델인 BERT를 사용하기로 함.

BERT는 NLP 11개의 테스트에 최고 성능을 기록(2020)했고 그 중 감정분석도 포함 되어있음.

Tasks	Dev Set							
	MNLI-m (Acc)	QNLI (Acc)	MRPC (Acc)	SST-2 (Acc)	SQuAD (F1)			
BERT _{BASE}	84.4	88.4	86.7	92.7	88.5	Fine-tuning approach		
No NSP	83.9	84.9	86.5	92.6	87.9			
LTR & No NSP	82.1	84.3	77.5	92.1	77.8			
+ BiLSTM	82.1	84.1	75.7	91.6	84.9			
						BERT _{LARGE}		
						96.6		
						92.8		
						BERT _{BASE}		
						96.4		
						92.4		

1. Word Piece 임베딩

Out of vocabulary(OOV) 문제를 해결함. + 형태소 분석을 할 필요없음

경찰, ##차, ##관, ##복 -> vocab에 없는 단어들에 대해서도 좋은 성능을 보임

* **한국어의 어미, 조사 등 학습하기 매우 용이함.**

2. 양방향 인코더

기존 NLP 관련 모델들은 좌에서 우로 입력 데이터를 학습시켰다. (앞에 단어를 통해 뒤에 단어를 예측)

하지만 버트는 양방향으로 학습시킴 (랜덤 추출하여 단어 하나를 비워 놓고 <MASK> 학습시킴)

* **단어 순서가 막 섞여도 말이 되는 한국어의 단점을 보완함**

3. 언어 모델링 데이터

총 33억 단어의 거대한 말뭉치를 이용하여 학습

***학습데이터가 많이 없어도 좋은 성능을 보임**

학습시간이 오래 걸리는 단점이 있지만

비교적 가벼운 BASE 모델을 사용하고 T3Q 클라우드 플랫폼(A100GPU)을 사용하여 단점을 보완.

문장 이진 분류 모델링

Transformer 패키지

Keras 사용

위키피디아에서 사전 학습시킨 모델 다운로드

bert-base-multilingual-cased 사용 (104 languages , 12-layer , 768-hidden , 12-head, 110M parameters)

- Config
- Checkpoint
- Vocab.txt

Bert_tokenizer를

Vocab 기준으로 임베딩 : input token

두 개의 문장을 구분하는 : segment input

Self-Attention 의 입력 위치를 나타내는 : position input

* 한국어의 경우에는 bert input을 제외하고 모두다 0을 사용함.

사전학습된 모델을 로드

BERT 모델 레이어 과정

1. 768차원으로 token , segment 임베딩
2. 12개의 셀프 어텐션 레이어
3. NSP , MLM

-----사전 학습 레이어-----

4. Dropout

5. Output : Dense(1, sigmoid)

5.BERT

테스트 정확도 : 80.01%

	Precision	recall	F1-score	support
0	0.74	0.73	0.74	1027
1	0.83	0.83	0.83	1549
accuracy			0.80	2576
macro avg	0.78	0.78	0.78	2576
Weighted avg	0.79	0.79	0.79	2576

사전 기반 감정 분류

Lexicon-based Approach

장점 : 수동으로 감정 사전을 구축하고 나면 적용하기 쉽다는 장점이 있다.

단점 : 도메인에 따라 사용하는 어휘가 달라지고 감정 점수도 달라지기 때문에 모든 도메인에 적용할 수 있는 감정사전을 구축하는 것이 매우 어렵다.

(1). 매뉴얼 기반 방법론

- 모든 단어에 대한 감성 사전을 수동으로 구축하는 방법
- 문서의 각 단어가 가지는 감정의 정도를 레이블링 한 것

(2). 사전 기반 방법론

- 매 분석마다 새로운 사전을 쓰는 것이 아니라 잘 구축되어 있는 외부 사전을 차용한다.
- 도메인 확장성이 없다.

ex) 영화리뷰에서 "줄리다" 는 부정
침대상품에서 "줄리다" 는 긍정 이런 경우 처리하기 어렵다.

(3) 말뭉치 기반 방법론

- 해당 말뭉치에 맞는 적절한 감성 어휘를 재구축하는 말뭉치 기반 접근 방법.
- 도메인 의존성을 극복할 수 있지만 좋은 사전 구축을 위해서 많은 데이터(거대 말뭉치)를 필요로 한다는 조건을 가지고 있다.

사전 기반 분류 모델

가사 도메인에 적합한 말뭉치 기반 감정 단어 사전 생성

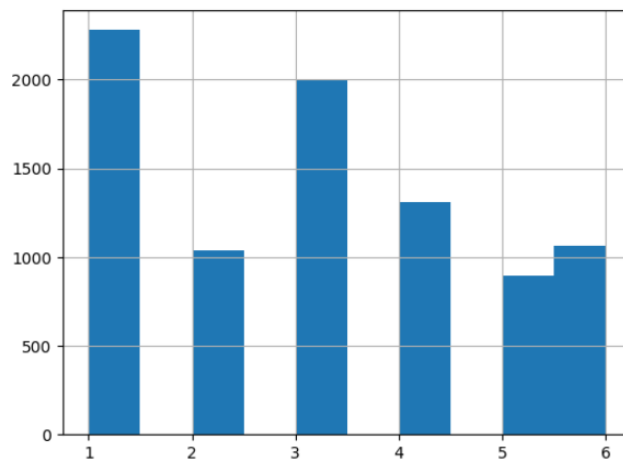
8586개의 문장을 6개의 감정으로 빈도 수 별로 단어 사전 생성

형태소 분석기는 : 코모란 형태소 분석기를 사용함.

단어사전에는 : **명사 : 일반명사, 고유명사, 의존명사**
형용사, 동사, 일반부사, 감탄사, 외국어만 포함시킴

동사와 부사에는 "+다"를 붙여서 사전에 저장함. *명사와 동사/부사를 구분 짓기 위함
(ex 사과, 사과하다)

	thrill	love	enjoy	sad	angry	empty
생각	0.56	0.04	0.11	0.12	0.09	0.09
눈	0.58	0.09	0.12	0.08	0.06	0.07
갈다	0.36	0.1	0.14	0.19	0.09	0.11
마음	0.47	0.11	0.06	0.26	0.01	0.09
되다	0.27	0.14	0.27	0.14	0.11	0.07
이별		0.02	0.04	0.95		
눈물			0.02	0.94		0.03
불안				0.92		0.08
슬프다		0.06		0.89	0.02	0.04
아픔		0.04		0.88		0.08



떨림	2279
사랑	1037
즐거움	2004
슬픔	1309
분노	896
공허	1061

감정 불용어 제거 (ex : 있다, 하다, 나다, 것, 너, 다...)

*있다는 감정과 아무런 상관이 없지만 제거하지 않는다면 감정 점수에 영향을 줌

사전 기반 분류 모델

형태소 분석기 비교

사전 기반 감정 분류 모델링을 위한 형태소 분석기는 정규화가 잘 되는 형태소 분석기를 사용하는 것이 좋다 형태는 다르지만 같은 뜻을 가지고 있는 단어들을 얼마만큼 하나로 묶어 주는 게 핵심이고 그 기준을 가지고 형태소 분석기를 선별했다.

정규화 성능 비교

문장의 수가 많아질 수록 같은 단어로 정규화된 단어 수를 표현

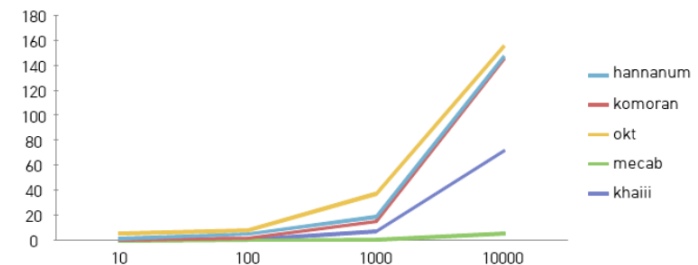
	1차(1000개)	2차(5000개)	3차(10000개)	4차(20000개)
꼬꼬마	597	915	1058	1232
코모란	566	876	999	자바메모리 에러
Okt	558	877	1004	1170
한나눔	558	877	1004	1170

*자바 메모리 에러 같은 경우 옵션을 줘서 더 할당해 줄 수 있지만 근본적인 해결책은 아님

속도 성능 비교

꼬꼬마 형태소 분석기는 속도가 너무 느림 제외

한나눔,코모란,okt 는 속도에서 비슷한 성능을 나타내고 Mecab의 경우 압도적으로 처리속도가 빠름



한국어 형태소 분석 블로그 자료 참고

사전 기반 분류 모델

형태소 분석기 비교

가사 문장 테스트

코모란

테스트 문장 :

'사랑을 잃은 이마음을'

'후회는 없잖아 그렇게 이유 없이 사랑했잖아'

'재미있게 듣던 말 이 책 중에 있으니'

'이렇게 지껄여도 아무 의미 없다'

'가만히 있어도 내 생각은 하는지'

```
[('사랑', 'NNG'), ('을', 'JKO'), ('잃', 'VV'), ('은', 'ETM'), ('이', 'MM'), ('마음', 'NNP'), ('을', 'JKO')]
[('후회', 'NNP'), ('는', 'JX'), ('없', 'VA'), ('잖아', 'EC'), ('그렇게', 'VA'), ('게', 'EC'), ('이유', 'NNG'), ('없이', 'MAG'), ('사랑', 'NNG'), ('하', 'XSV'), ('았', 'EP'), ('잖아', 'EC')]
[('재미있', 'VA'), ('게', 'EC'), ('들', 'VV'), ('던', 'ETM'), ('말', 'NNG'), ('이', 'MM'), ('책', 'NNG'), ('중', 'NNB'), ('에', 'JKB'), ('있', 'VV'), ('으니', 'EC')]
[('이렇게', 'VA'), ('게', 'EC'), ('지껄여', 'VV'), ('어도', 'EC'), ('아무', 'MM'), ('의미', 'NNG'), ('없', 'VA'), ('다', 'EC')]
[('가만히', 'MAG'), ('있', 'VV'), ('어도', 'EC'), ('내', 'NP'), ('생각', 'NNG'), ('은', 'JX'), ('하', 'VV'), ('는지', 'EC')]
```

Okt (norm =True , stem = True) * **norm** 은 정규화 , **stem** 은 어간추출 (ex 해야지 -> 하다)

```
[('사랑', 'Noun'), ('을', 'Josa'), ('잃다', 'Verb'), ('이', 'Determiner'), ('마음', 'Noun'), ('을', 'Josa')]
[('후회', 'Noun'), ('는', 'Josa'), ('없다', 'Adjective'), ('그렇게', 'Adverb'), ('이유', 'Noun'), ('없이', 'Adverb'), ('사랑', 'Noun'), ('하', 'Verb')]
[('재미있다', 'Adjective'), ('들다', 'Verb'), ('말', 'Noun'), ('이', 'Noun'), ('책', 'Noun'), ('중', 'Noun'), ('에', 'Josa'), ('있다', 'Adjective')]
[('이렇게', 'Adverb'), ('지껄이다', 'Verb'), ('아무', 'Noun'), ('의미', 'Noun'), ('없다', 'Adjective')]
[('가만히', 'Adverb'), ('있다', 'Adjective'), ('내', 'Noun'), ('생각', 'Noun'), ('은', 'Josa'), ('하다', 'Verb')]
```

한나눔의 경우는 "이마음" 이라고 분석한 부분에 있어서 "마음", "이마음", "맘" 이렇게 같은 의미의 단어를 정규화하는 부분의 성능이 좋지않아 제외했다.

Komoran 과 Okt 를 두고 고민을 많이 했는데 (둘다 속도가 비슷)
정규화 측면에서 명사는 Okt 의 성능이 좋았고 , 동사는 코모란의 성능이 좋았다.

사실 두 형태소 분석기의 장/단점이 있지만
직관적으로 보았을 때 보기 편한 방법으로 코모란 형태소에 부사와 , 동사에 '+다'를 붙여주는 방식을 채택하였다.

사전 기반 분류 모델

코모란 형태소 분석기 결과

```
[('사랑', 'NNG'), ('을', 'JKO'), ('잃', 'VV'), ('은', 'ETM'), ('이', 'MM'), ('마음', 'NNP'), ('을', 'JKO')]
[('후회', 'NNP'), ('는', 'JX'), ('없', 'VA'), ('잖아', 'EC'), ('그렇', 'VA'), ('게', 'EC'), ('이유', 'NNG'), ('없이', 'MAG'), ('사랑', 'NNG'),
('하', 'XSV'), ('있', 'EP'), ('잖아', 'EC')]
[('재미있', 'VA'), ('게', 'EC'), ('들', 'VV'), ('던', 'ETM'), ('말', 'NNG'), ('이', 'MM'), ('책', 'NNG'), ('중', 'NNB'), ('에', 'JKB'), ('있',
'VV'), ('으니', 'EC')]
[('이렇', 'VA'), ('게', 'EC'), ('지절이', 'VV'), ('어도', 'EC'), ('아무', 'MM'), ('의미', 'NNG'), ('없', 'VA'), ('다', 'EC')]
[('가만히', 'MAG'), ('있', 'VV'), ('어도', 'EC'), ('내', 'NP'), ('생각', 'NNG'), ('은', 'JX'), ('하', 'VV'), ('는지', 'EC')]
```

테스트 문장 :

'사랑을 잃은 이마음을'

'후회는 없잖아 그렇게 이유 없이 사랑했잖아'

'재미있게 듣던 말 이 책 중에 있으니'

'이렇게 지절여도 아무 의미 없다'

'가만히 있어도 내 생각은 하는지'

형태소 분석
필요 품사 필터링

```
[ '사랑', '잃다', '마음' ]
[ '후회', '없다', '그렇다', '이유', '없이', '사랑' ]
[ '재미있다', '듣다', '말', '책', '중', '있다' ]
[ '이렇다', '지절이다', '의미', '없다' ]
[ '가만히', '있다', '생각', '하다' ]
```

명사 : 일반명사, 고유명사, 의존명사
형용사, 동사, 일반부사, 감탄사, 외국어만 포함시킴

동사 & 부사에는 '+다'를 붙여서 형태소 분석

결과를 사전에 저장함.

사전 기반 분류 모델 테스트

테스트는 긍/부정에 따른 사전을 따로 분리하지 않고 모든 감정에 대해서 진행했습니다.

* 포함되는 단어에 감정 점수만 고려한 결과 (사전으로만 감정을 분류했을 경우)

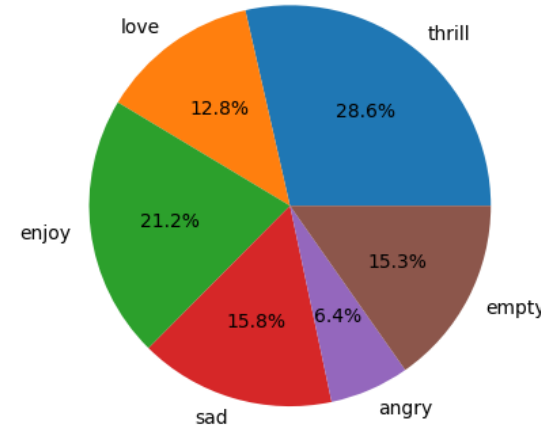
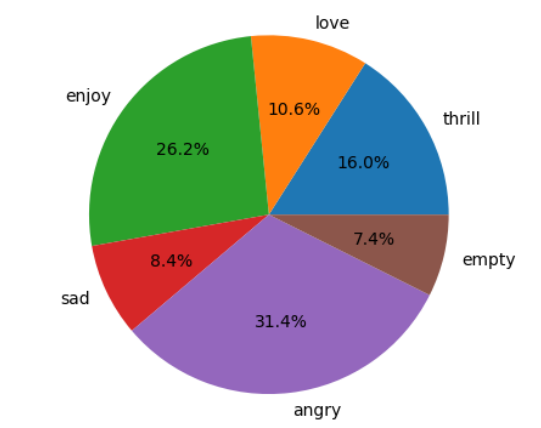
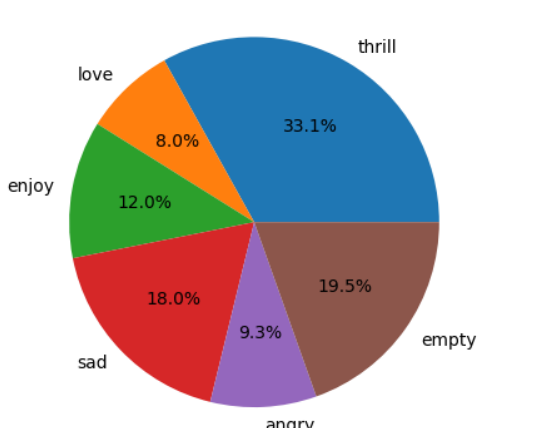
Best case :

지루할 틈 없이 너만 기다리는 On your pet	아이처럼 한번 웃어봐 옛날처럼 딱 한번만	흐르는 빗물이 이별을 말해
예측 결과 : '떨림' 정답 : '떨림'	예측 결과 : '즐거움' 정답 : '즐거움 '	예측 결과 : '슬픔' 정답 : '슬픔'

사전 기반 분류 모델 테스트

테스트는 긍/부정에 따른 사전을 따로 분리하지 않고 모든 감정에 대해서 진행했습니다.

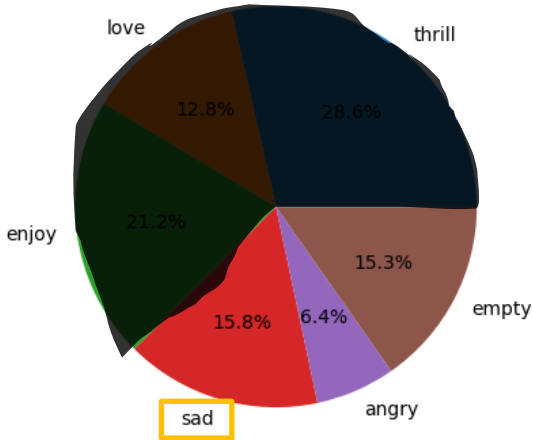
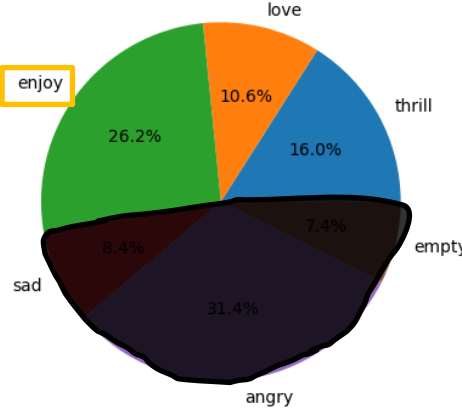
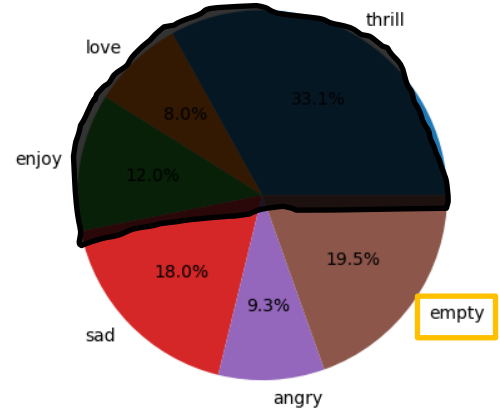
Worst case :

시간의 흐름은 너와 나는 반대로	때론 성공만을 바라던 내게 질책아닌 해결책을	거짓없는 그대의 미소가 너무 초라해 보여																																										
예측 결과 : '떨림' 정답 : '슬픔'	예측 결과 : '분노' 정답 : '즐거움'	예측 결과 : '떨림' 정답 : '공허'																																										
 <table><thead><tr><th>Emotion</th><th>Percentage</th></tr></thead><tbody><tr><td>thrill</td><td>28.6%</td></tr><tr><td>love</td><td>12.8%</td></tr><tr><td>enjoy</td><td>21.2%</td></tr><tr><td>sad</td><td>15.8%</td></tr><tr><td>angry</td><td>6.4%</td></tr><tr><td>empty</td><td>15.3%</td></tr></tbody></table>	Emotion	Percentage	thrill	28.6%	love	12.8%	enjoy	21.2%	sad	15.8%	angry	6.4%	empty	15.3%	 <table><thead><tr><th>Emotion</th><th>Percentage</th></tr></thead><tbody><tr><td>angry</td><td>31.4%</td></tr><tr><td>love</td><td>10.6%</td></tr><tr><td>thrill</td><td>16.0%</td></tr><tr><td>empty</td><td>7.4%</td></tr><tr><td>sad</td><td>8.4%</td></tr><tr><td>enjoy</td><td>26.2%</td></tr></tbody></table>	Emotion	Percentage	angry	31.4%	love	10.6%	thrill	16.0%	empty	7.4%	sad	8.4%	enjoy	26.2%	 <table><thead><tr><th>Emotion</th><th>Percentage</th></tr></thead><tbody><tr><td>thrill</td><td>33.1%</td></tr><tr><td>love</td><td>8.0%</td></tr><tr><td>enjoy</td><td>12.0%</td></tr><tr><td>sad</td><td>18.0%</td></tr><tr><td>angry</td><td>9.3%</td></tr><tr><td>empty</td><td>19.5%</td></tr></tbody></table>	Emotion	Percentage	thrill	33.1%	love	8.0%	enjoy	12.0%	sad	18.0%	angry	9.3%	empty	19.5%
Emotion	Percentage																																											
thrill	28.6%																																											
love	12.8%																																											
enjoy	21.2%																																											
sad	15.8%																																											
angry	6.4%																																											
empty	15.3%																																											
Emotion	Percentage																																											
angry	31.4%																																											
love	10.6%																																											
thrill	16.0%																																											
empty	7.4%																																											
sad	8.4%																																											
enjoy	26.2%																																											
Emotion	Percentage																																											
thrill	33.1%																																											
love	8.0%																																											
enjoy	12.0%																																											
sad	18.0%																																											
angry	9.3%																																											
empty	19.5%																																											

문맥을 보지 않고 단어 자체로만 판단한다면 이런 오분류를 범할 수 있음.

사전 기반 분류 모델 테스트

이진 분류(긍/부정)를 정확하게 분류 했더라면?

시간의 흐름은 너와 나는 반대로	때론 성공만을 바라던 내게 질책아닌 해결책을	거짓없는 그대의 미소가 너무 초라해 보여
예측 결과 : '떨림' 정답 : '슬픔'	예측 결과 : '분노' 정답 : '즐거움'	예측 결과 : '떨림' 정답 : '공허'
		

하이브리드 모델에 전 과정인

긍/부정 분류가 잘 되었다고 가정하면 제대로 분류한다고도 볼 수 있다. 그래서 **신경망 기반 이진분류**가 더 중요하다.

한국어 가사 감정 분석 흐름도

문장 분할

감정 사전 대입

BERT 모델

긍정 : 슬픔 , 분노 , 공허 제외
부정 : 떨림 , 사랑 , 즐거움 제외

A 문장

감정 점수
스코어링

긍/부정 분류

긍정 or 부정
감정만 남김

B 문장

감정 점수
스코어링

긍/부정 분류

긍정 or 부정
감정만 남김

C 문장

감정 점수
스코어링

긍/부정 분류

긍정 or 부정
감정만 남김

D 문장

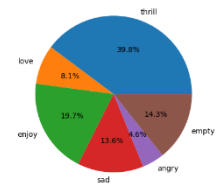
감정 점수
스코어링

긍/부정 분류

긍정 or 부정
감정만 남김

가사

합산



한국어 가사 감정 분석 결과

일시	학습데이터 양	BERT 정확도	감정 단어 수	감정 라벨 수
21/05/20	8789	80.01%	2822	6개

	가수	제목	장르	1위 감정	2위 감정
1	장범준	흔들리는 꽃들 속에서 니 샴푸향이 느껴진거야	발라드	떨림	즐거움
2	브레이브 걸스	롤린	아이돌/댄스	떨림	즐거움
3	SG워너비	살다가	발라드	슬픔	공허
4	IU	내 손을 잡아	발라드	떨림	즐거움
5	IU / 하이포	봄 사랑 벚꽃 말고	발라드	떨림	즐거움
6	윤종신	좋니	발라드	떨림	즐거움
7	머쉬베놈	VVS	힙합	즐거움	떨림
8	창모	METEOR	힙합	즐거움	떨림
9	지코	아무노래	힙합	즐거움	떨림
10	산이	니가 불행했음 좋겠다	힙합	분노	즐거움
11	10cm	스토커	어쿠스틱	슬픔	즐거움
12	박구윤	뿐이고	트로트	떨림	즐거움
13	홍진영	엄지 척	트로트	사랑	떨림
14	god	거짓말	아이돌/발라드	슬픔	분노
15	성시경	내게 오는 길	발라드	떨림	사랑
16	버즈	가시	발라드	슬픔	공허
17	이기찬	미인	발라드	슬픔	공허
18	트와이스	KNOCK KNOCK	아이돌/댄스	즐거움	떨림
19	트와이스	Cheer up	아이돌/댄스	떨림	즐거움
20	방탄소년단	피 땀 눈물	아이돌/댄스	즐거움	슬픔
21	방탄소년단	작은 것들을 위한 시	아이돌/댄스	즐거움	떨림

[테스트 결과](#)
[원노트 링크](#)

영어 가사 감정 분류

한국어 처럼 영어도 마찬가지로 가사도메인에 대한 데이터셋을 구축하여 분석을 진행하는 것이 가장 좋다고 생각하지만
영어에 대한 라벨링을 진행하는 것이 시간 소요가 많고 정확도가 많이 떨어질 수 있다고 생각하여
영어에 대한 감정을 나타내는 다른 TEXT 데이터 셋을 활용하여 감정 분류를 진행

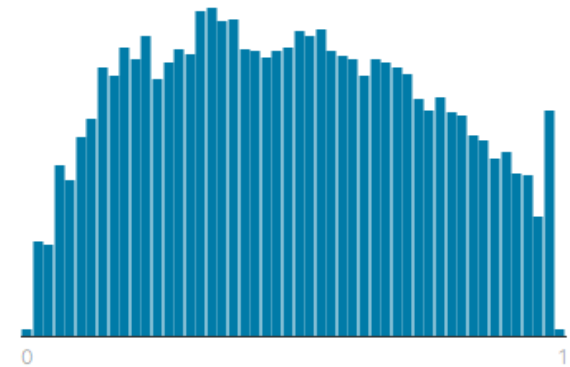
영어 가사 감정 데이터 셋

150K Lyrics Labeled with Spotify Valence

출처 : <https://www.kaggle.com/edenbd/150k-lyrics-labeled-with-spotify-valence/code>

0	Elijah Blake	No, no I ain't ever trapped out the bando But oh Lord, don't get me wrong I know a couple niggas ...	Everyday	0.626
1	Elijah Blake	The drinks go down and smoke goes up, I feel myself, got to let go My cares get lost up in that cro...	Live Till We Die	0.63
2	Elijah Blake	She don't live on planet Earth no more She found love on Venus, that's her word Said she needed sp...	The Otherside	0.24

총 데이터 갯 수: 15만



세부 감정에 대한 라벨링은 아니지만 가사 도메인이라는 점에서 좋은 데이터 셋
전 처리가 어느정도 이미 되어있는 cleaned data

영어 가사 감정 분석 방법론

1 . 감정 예측 + 가사 전체 분위기

딥러닝 모델을 통해서 문장별 감정을 예측하고 전체적인 분위기 긍정/부정에 대한 가중치를 입히는 방법.

평서문 감정 & 트위터 감정으로 예측한 감정 분류 결과에 가사 전체에 대한 긍정/부정 결과를(가사분위기) 가중치를 추가로 입혀줌.

Ex) 가사 전체 감정 결과가 긍정 70% 부정 30% 이 나왔다면 , 긍정 감정에 추가 가중치 적용

가사 전체 분위기 (긍정적 가사 / 부정적 가사) 를 파악하기 위해서

150K Lyrics Labeled with Spotify Valence 데이터 셋을 가지고 BERT 이진 분류 시도

0.5를 기준으로 긍정/부정 치환 *0.5 이하는 0 , 0.5 이상은 1

그러나 데이터 셋이 많음(15만건)에도 불구하고 정확도는 63% 이상 올라가지 않음.

*가사만 가지고 긍정/부정을 판단하는 것조차 어려움 , Kaggle 노트북에서도 시도한 사람이 없음.

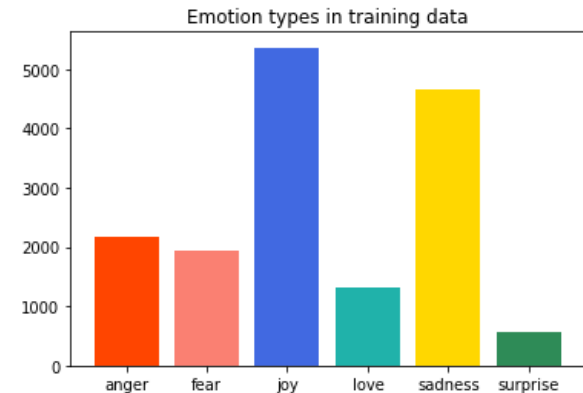
좀 더 연구해볼 가치는 있었지만 시간이 길어질 것 같아서 포기

영어 가사 감정 데이터 셋

Emotion dataset for NLP : 평서문 감정 데이터

출처:<https://www.kaggle.com/praveengovi/emotions-dataset-for-nlp>

```
im feeling rather rotten so im not very ambitious right now;sadness  
im updating my blog because i feel shitty;sadness  
i never make her separate from me because i don t ever want her to feel  
i left with my bouquet of red and yellow tulips under my arm feeling sli  
i was feeling a little vain when i did this one;sadness  
i cant walk into a shop anywhere where i do not feel uncomfortable;fear  
i felt anger when at the end of a telephone call;anger
```



데이터 불균형

감정 종류

총 6가지로 놀람, 사랑, 즐거움, 슬픔, 화남, 두려움이 있다.

평서문 데이터로 유연한 문장이 없음 직접적인 문장이 대부분
가사 데이터는 비유적 표현이나 돌려말하는 부분이 많은 것에 비해서는 단점

영어 가사 감정 분석 방법론

2) 문장 긍/부정 분류기 + 단어 사전 점수 (기존 한국어 가사 감정 분류 방식)

평서문 데이터를 가지고 감정들을 긍/부정으로 치환하여 학습데이터 셋팅

{'surprise' : 1 , 'love' : 1 , 'joy' : 1 , 'sadness' : 0 , 'anger' : 0 , 'fear' : 0}

위와 동일한 모델링을 진행하고

Output layer 수정 : Dense(1,activation = 'sigmoid')

손실함수 및 평가함수 변경

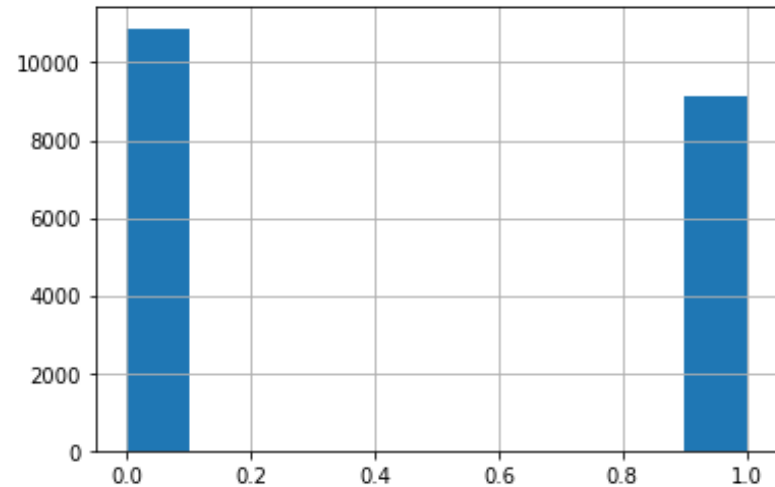
Loss : BinaryCrossentropy

Metric : Accuracy

	precision	recall	F1-score	support
0	0.99	0.97	0.98	3264
1	0.96	0.99	0.98	2736

Accuracy : 98%

가사데이터에 테스트



영어 가사 감정 분석 방법론

평서문 데이터 다중 분류 모델링

Bert Base 모델 Distilbert 사용

* BERT를 40%로 줄이고 60% 빠르게 연산하면서
97%의 성능을 유지함.

학습데이터 : 14000

검증데이터 : 7000

MAX_LEN = 70

Input : input_ids , input_mask

Layers :

GlobalMaxPool1Dd()(embeddings)

Dense(128 , activation = 'relu')

Dropout(0.1)

Dense(32, activation = 'relu')

#output

Dense(6, activation = 'sigmoid')

Compile :

Optimizer : Adam

Loss : CategoricalCrossentropy

Metric : categoricalAccuracy

Fit :

Epoch : 1

Batch_size = 36

Accuracy	0.9259
F1 score	0.8842

Classification Report:

	Precision	Recall	F1-score	Support
0	0.92	0.97	0.94	664
1	0.97	0.96	0.97	586
2	0.90	0.94	0.92	263
3	0.94	0.85	0.89	247
4	0.90	0.77	0.83	185
5	0.68	0.83	0.75	54

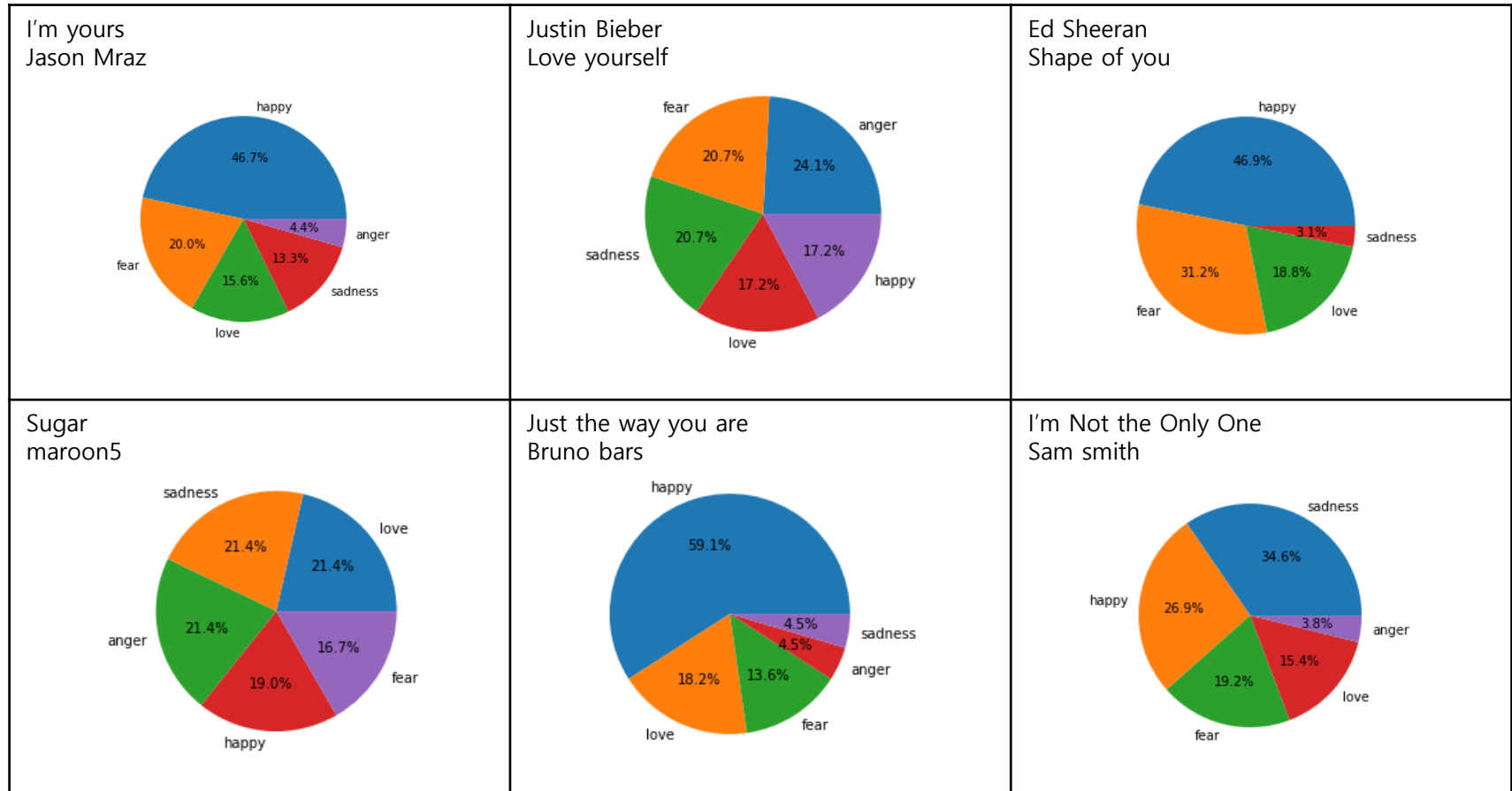
Confusion Matrix:

[[642	4	3	0	15	0]
[4	563	10	6	0	3]
[4	5	248	5	1	0]
[0	6	13	210	0	18]
[39	2	1	0	143	0]
[6	0	0	3	0	45]]]

정제가 잘 되어있는 데이터 셋이고 평서문데이터는 특정한 규칙이 존재하기 때문에 학습이 잘 된듯합니다.

영어 가사 감정 분석 방법론

평서문 데이터 셋 가사 감정 분류 - 가사 데이터에 테스트



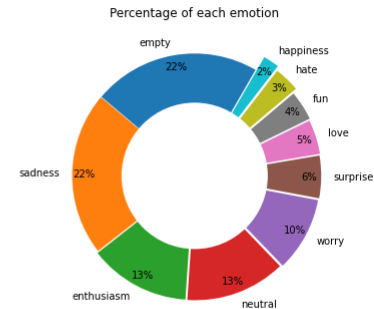
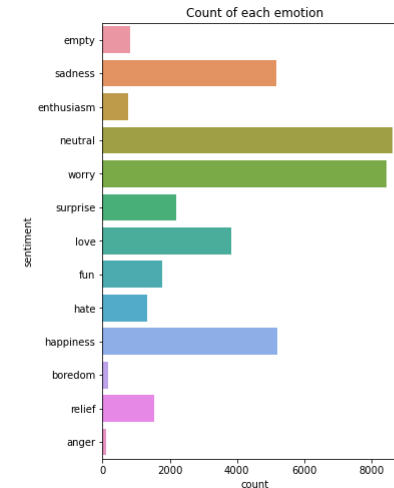
결과는 나쁘지 않지만 fear 감정이 일정 비율로 모두 포함 되어있음.

영어 가사 감정 데이터 셋

Emotion Detection from Text : 트위터 감정 데이터

출처 : <https://www.kaggle.com/pashupatigupta/emotion-detection-from-text>

	tweet_id	sentiment	content
0	1956967341	empty	@tiffanylue i know i was listenin to bad habi...
1	1956967666	sadness	Layin n bed with a headache ughhhh...waitin o...
2	1956967696	sadness	Funeral ceremony...gloomy friday...
3	1956967789	enthusiasm	wants to hang out with friends SOON!
4	1956968416	neutral	@dannycastillo We want to trade with someone w...



감정 종류

총 13가지로 놀람,사랑,재밌는,열정,걱정,무감정,슬픔,화남, 싫은,행복,지루한,편안함,공허함이 있다

총 데이터 갯 수: 40000개의 트윗 데이터

마찬가지로 불균형이 심함

- 트윗 데이터의 경우 오타자가 심하고 문법적인 오류가 많이 있다.
- 인터넷 문자 용어(줄임말)이 많이 사용됨
- you -> u
- and -> n
- yeeeeeahhh~!!
- oooooooooohhh..

영어 가사 감정 분석 방법론

3) 트위터 + 평서문 데이터 셋을 활용한 다중 분류 모델 (불균형 문제 해결)

평서문 감정 분류 모델에 데이터 셋만 교체

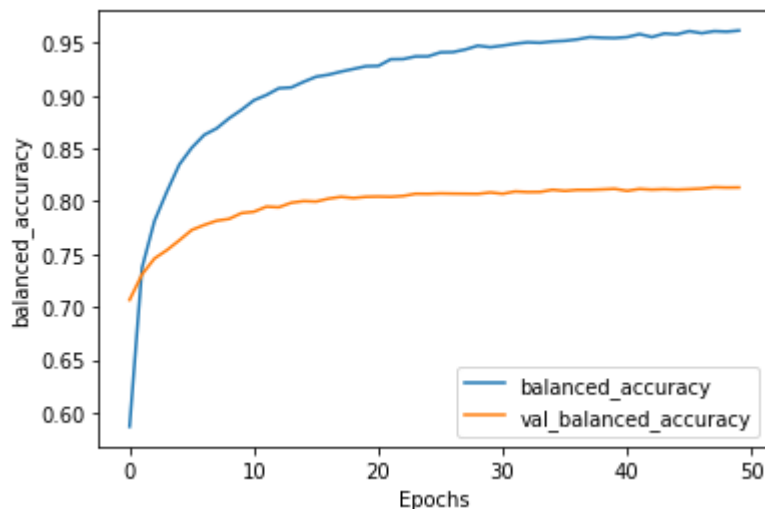
학습데이터 : 20300개

검증데이터 : 8700개

Epochs = 50

Batch_size = 64

	Tweet	Natural	Total
surprise	2000	1000	3000
Love	3000	2000	5000
Happy	3000	3000	6000
Sadness	3000	3000	6000
Anger	2000	3000	5000
Fear	1000	3000	4000



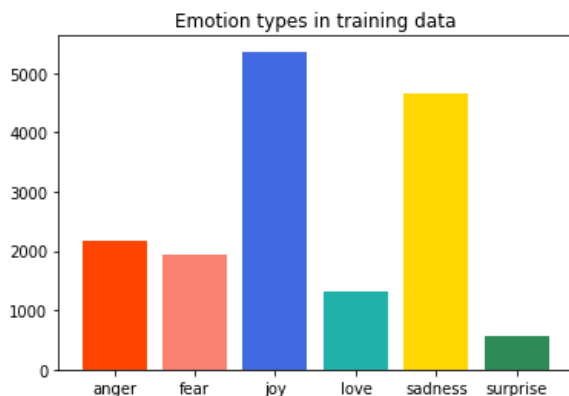
	precision	recall	f1-score	support
0	0.70	0.71	0.70	900
1	0.79	0.81	0.80	1500
2	0.81	0.76	0.78	1800
3	0.81	0.81	0.81	1800
4	0.86	0.90	0.88	1500
5	0.88	0.88	0.88	1200
accuracy			0.81	8700
macro avg	0.81	0.81	0.81	8700
weighted avg	0.81	0.81	0.81	8700

영어 가사 감정 분석 방법론

사전 기반 감정 분석

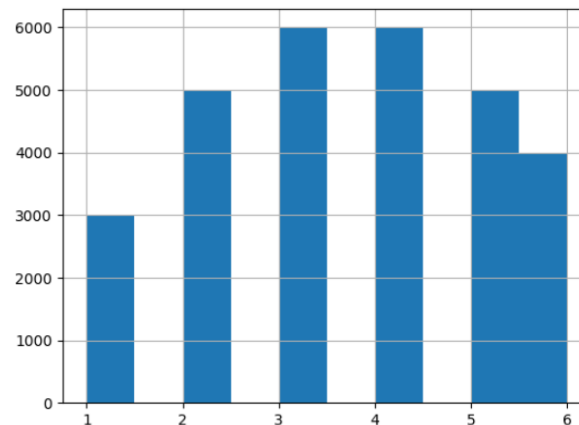
- 빈도 수 기반으로 사전을 구축하는 말뭉치 사전 구축 방법은 데이터 불균형 문제가 아주 치명적임.
- 전체적인 볼륨이 큰 라벨이 대부분의 점수를 차지하는 문제 발생

*데이터 불균형 발생



트위터 데이터

Under & over 샘플링



Joy와 sadness의 데이터가 상대적으로 많기 때문에 대부분의 점수가 joy와 sadness의 많은 비중을 차지함

심지어 "love" 단어도 joy의 비중이 더 큼

데이터 불균형 문제 해결

영어 가사 감정 분석 방법론

감정 단어 사전 생성

전처리(pre-processing)

1. 특수문자 제거(정규표현식)

- http 포함된 단어 제거
- @ 포함된 단어 제거
- 영어 숫자가 아닌 글자 제거
- 소문자 변환 / 앞뒤 공백 제거

2. 불용어 제거 stopwords.words("English")

- nltk.corpus 에 포함되어있는 stopwords 를 사용함.
- 100여개 이상의 영어 단어들을 불용어로 패키지 내에서 미리 정의하고 있음.

3. 너무 짧은 문장 제거 15개 이하의 글자 수 데이터 제거

4. 토큰나이징(word_tokenize)

- 문장을 단어단위로 토큰화

5. 형태소 분석(nltk.pos_tag)

- 토큰화 된 단어들의 형태소를 판별함.

6. 정규화 (WordNetLemmatizer)

- 명사의 복수형을 단수로 바꿔주고
- 불규칙/규칙 동사들을 전부 동사 원형으로 바꿔줌.

7. 사전에 정제

- 동사(VB) , 형용사(JJ) , 부사(RB), 명사(NN)만 단어 사전에 추가

Word	Lemmatized
Dies	Die
Watched	Watch
Has	Have

처리 전 :

@tiffanylue i know i was listenin to bad habit earlier and i started freakin at his

처리 후:

I Know I be listenin bad habit earlier I start freakin

영어 가사 감정 분석 방법론

감정 단어 사전 결과

	surprise	love	happy	sadness	anger	fear
day	0.08	0.32	0.25	0.18	0.09	0.07
think	0.14	0.14	0.2	0.19	0.19	0.14
really	0.11	0.18	0.21	0.18	0.2	0.11
time	0.1	0.15	0.24	0.2	0.17	0.14
look	0.14	0.19	0.21	0.19	0.13	0.13
still	0.13	0.14	0.17	0.22	0.16	0.18
love	0.07	0.6	0.16	0.07	0.06	0.05
good	0.08	0.26	0.35	0.17	0.09	0.06
funny	0.57	0.11	0.19	0.09	0.05	
find	0.18	0.15	0.19	0.16	0.15	0.17
new	0.14	0.21	0.25	0.18	0.1	0.11
amaze	0.49	0.23	0.27		0.02	

총 단어수 : 5095개

영어 가사 감정 분석 테스트

일시	테스트 모델	학습데이터 양	BERT 정확도	감정 단어 수	감정 라벨 수
21/05/31	(1) 긍/부정 + 사전 기반 모델	21,459개	98% (binary)	5095	6개
	(2) 감정 분류 모델	21,459개	92% (multi class)	X	6개
21/06/01	(3) 감정 분류 모델 밸런스 (트위터데이터 추가 균형있는 데이터)	29,900개 (tweet + natural)	81% (multi class)	X	6개

가수	제목	1번	3번				
Justin Bieber	Love Yourself	분노	슬픔	Mark Ronson	Uptown Funk	사랑	놀람&두려움
Ed Sheeran	Shape of You	사랑	두려움	Owl City	Fireflies	분노	두려움
Charlie Puth	See You Again	행복	사랑	Taylor Swift	Black Space	행복	두려움
Jason Mraz	I'm Yours	사랑	행복	Britney Spears	Toxic	분노	두려움
Adele	Someone Like You	사랑	행복	Michael Jackson	Billie Jean	사랑	두려움
Maroon5	Sugar	사랑	슬픔&사랑	Queen	We Are The Champions	슬픔	행복
Adam Levine	Lost Stars	분노	슬픔	Muse	Time Is Running Out	슬픔	분노
Bruno Mars	Just the Way You Are	행복	행복&사랑	Rihanna	Diamonds	행복	행복
Sam Smith	I'm Not the Only One	분노	슬픔	Eminem	Lose yourself	슬픔	슬픔&분노
Coldplay	Viva La Vida	슬픔	두려움	Eminem	Not Afraid	행복	분노
				Eminem	Without Me	분노	분노

[테스트 결과 원노트 링크](#)

추가 방안

한국어 감정 추가 방안 :

즐거움	신나는	공허	외로움
	열정		그리움
	행복한		

감정을 추가하면 세분화된 감정으로 분류 할 수 있다는 장점이 있지만 감정 단어 사전 분석이 빈도 수를 기반으로 하기 때문에 즐거움의 빈도 수를 세 가지로 나누고 공허를 두 가지로 나누면 빈도 수가 그만큼 줄어 들 수 있고 점수 배정이 떨어 질 수 있다.

예를 들어 즐거움이 1위 감정이 었던 음원이 감정이 세분화 되면서 다른 감정이 1위로 올라갈 수 있는 경우가 발생한다. 그래서 충분한 라벨데이터가 균형 있게 존재하는 상황에서 다시 감정 단어 사전을 재 구축하고 분석해야 좋은 결과를 얻을 수 있다.

영어 가사 감정 분류 방법론 두 가지를 합산하는 방안 :

하이브리드 모델 + 다중 분류 모델

이렇게 되면 하이브리드 모델로 **단어적 감정**을 반영하고 다중 분류 모델로 **문맥적 감정**을 반영하여 더 정확한 결과를 얻을 수 있다고 생각한다.

결론 :

여러가지 데이터 셋으로 테스트하고 경험해본 결과 성능향상에 가장 큰 도움이 되는 것은 **가사 도메인으로 된 대량의 감정 데이터 셋이 필요하다.**

추가 과제

파일 및 코드정리

- 주석 및 docstring 작성
- 전처리 및 실행 함수 모듈화
- 데이터 셋만 교체하면 분석 결과 나오게끔 셋팅
* 추후 라벨링을 더 진행하거나 새로운 데이터 셋을 찾을 경우 대비
- NAS에 백업
- 팀즈 - 원 노트 목차 정리

DB 가사 데이터 점검

- 데이터 베이스에 있는 가사 데이터 점검(형식이 안 맞거나 가사데이터가 없는 경우)
- 가사 데이터 정규화 작업

분석 시간 측정

- 결과 분석 시간을 계산하여 실제 데이터베이스에 채워 넣을 때 예상 소요 시간 계산

기타 업무

MySQL DB 데이터 Python 으로 처리하고 데이터 추가

쿼리로 처리하기 힘들고 python dataframe으로 쉽게 처리할 수 있는 문제는 바로 DB에 접근해서 처리할 수 있도록 정리해 두었음.

* RENEW_music_artist_kr 테이블 데이터 추가

BPM 추출 자동화

BPM 추출 API 를 사용해서 추출하고 txt 파일에 정제(쉼표로 구분) 어느정도 완료되면 중간중간 추합해서 저장.

*** BPM 추출 에러 음원 해결해야함**

윈도우 WSL 환경 셋팅 메뉴얼화

업무 관리

업무 관리 경향

오전 9시 30분 ~ 오후 6시 30분 기준 근무 선호

오전 근무 : 개발 내용 정리(원 노트) + 업무 관리 갱신

오후 근무 : 개발 업무

퇴근 1시간 전 : 라벨링

업무 관리 방법

Teams 를 활용한 업무관리

업무관리 탭에서 작업 추가 및 갱신

- 중요 핵심내용 위주로 작성

자세한 업무 내용은 원노트에 기입(전문적인 내용)

시행착오 및 개선과정 원노트에 정리

*** 모든 내용 인과관계를 중점으로 작성**

감사합니다
