

MOBILE BANKING APP REVIEWS ANALYSING

1015 Process Presentation, team 4



TABLE OF CONTENTS

01

OBJECTIVES

주제와 목표 구체화

02

DATA

데이터의 수집 과정과
수집된 리뷰 데이터의 정보

03

SCORING

씨티은행 리뷰 데이터의
명사/문장 점수 매기기 예시

04

PROJECTS

문제제기와 앞으로의 계획

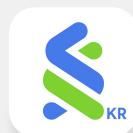
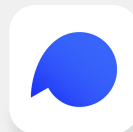
01

OBJECTIVES

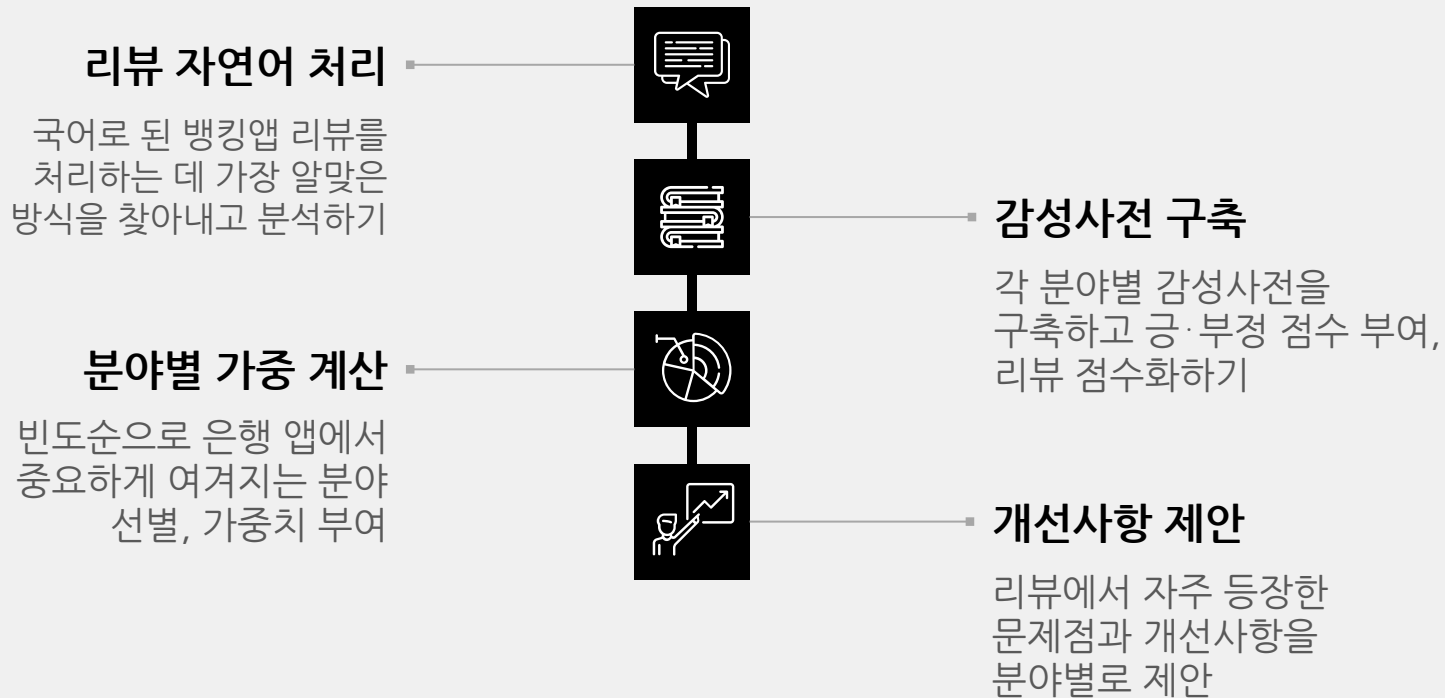
주제와 목표 구체화

모바일뱅킹 앱 리뷰

금융소비자연맹에서 선정한 은행 순위에서
지방은행을 제외한 상위 11개 은행의
모바일뱅킹 앱 리뷰를 분석



GOALS



DOMAINS – variable



기능 / 상품

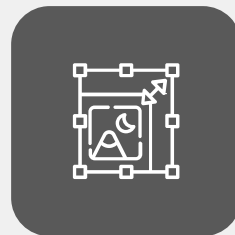
각 은행에서 제공하는
특별 기능이나 상품

Ex. 카카오뱅크의 모임통장 기능



기술적 문제

보안, 앱 실행 및 종료,
속도, 로그인 문제 등



디자인

UX, UI, 심미성, 직관성,
보기 불편한 메뉴 등


02


DATA


데이터의 수집 과정과
수집된 리뷰 데이터의 정보

DATA 수집 과정

데이터 수집처  Google Play

**Yunkyung kim**
★★★★★ 2021년 7월 16일
다른은행 모바일앱보다 보기에도 사용하기에도 모두 편해요.

**inwoo kong**
★★★★★ 2021년 9월 24일
최악입니다. 오류떠서 지우고 다시 깔았는데, 인증 무한루프입니다. 짜증나네요

**NEO GEO**
★★★★★ 2021년 9월 12일
간편한데 앱이 최적화가 잘안되어있는것같아요. 너무 무겁게 느낌.

11개 은행 App에서
총 193,715개의 리뷰 수집



Crawling
google-play-scraper

google-play-scraper 1.0.2

```
pip install google-play-scraper
```

PYPI에서
google-play-scraper
패키지 사용

TRIAL TIMELINE



TRIAL TIMELINE

Stopwords Removal

분석 및 점수계산에 큰
의미가 없는 불용어 제거

WordCloud& Histogram

모든 단어 및 가장 많이
나온 40개 단어에 대해
WordCloud 생성

Morpheme Analysis

komoran을 이용하여
형태소 분석&품사 분리

Score Calculation

점수계산 알고리즘
구상중

CODE & OUTPUTS

```
df = pd.DataFrame(columns=['userName', 'score', 'date', 'thumbsUpCount', 'comment'])
review_cnt = 1

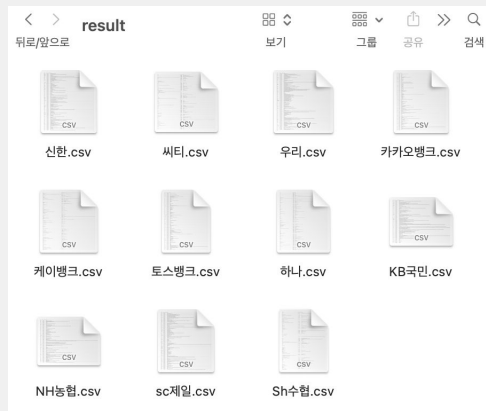
for review in reviews:

    userName = review['userName']
    score = review['score']
    date = review['at']
    thumbsUpCount = review['thumbsUpCount']
    comment = review['content']

    df = df.append({
        'userName': userName,
        'score': score,
        'date': date,
        'thumbsUpCount': thumbsUpCount,
        'comment': comment
    }, ignore_index=True)

print(review_cnt, "번째 리뷰 수집 완료")
review_cnt += 1
```

google-play-scraper를 이용한
crawling code 일부



전체 기간에 대해 각 모바일뱅킹 앱의
모든 리뷰를 csv 파일형태로 수집



2021_sent_token

1	편리성이 너무 떨어짐 앱 다루기가 힘들
2	저도 몇일 동안오류로 뜨네요.
3	설정 들어가면 아무것도 안되고 먹통이되요 그리고 원하는걸 하기어
4	앱 ui 진짜 너무 답답하고 불편해요 디자인은 말할것도 없고... 그리
5	최악입니다.
6	터치가 안먹혀요 ... 답답답답답답 속터짐 ㅎㅎ 언제 고쳐주시나요
7	차라리 앱을 없애는게 나을 것 같아요.

NLTK -> 문장별 분리



2021_morph_token

1	[('편리', 'NNG'), ('너무', 'MAG'), ('떨어지', 'VV'), ('다루', 'VV'), ('힘들', 'VV')]
2	[('일', 'NNP'), ('동안', 'NNG'), ('오류로', 'NNP'), ('뜨', 'VV')]
3	[('설정', 'NNG'), ('들어가', 'VV'), ('아무것도', 'NNG'), ('안', 'NNG'), ('먹', 'VV')]
4	[('진짜', 'MAG'), ('너무', 'MAG'), ('답답', 'XR'), ('불편', 'NNG'), ('디자', 'NNP')]
5	[('최악', 'NNG'), ('아', 'VCP')]
6	[('터치', 'NNP'), ('안', 'MAG'), ('먹히', 'VV'), ('답답', 'XR'), ('답답', 'X')]
7	[('차라리', 'MAG'), ('없애', 'VV'), ('나', 'VV'), ('같', 'VA')]

Komoran -> 형태소/품사 분리

star_1 ▾

star_2 ▾

star_3 ▾

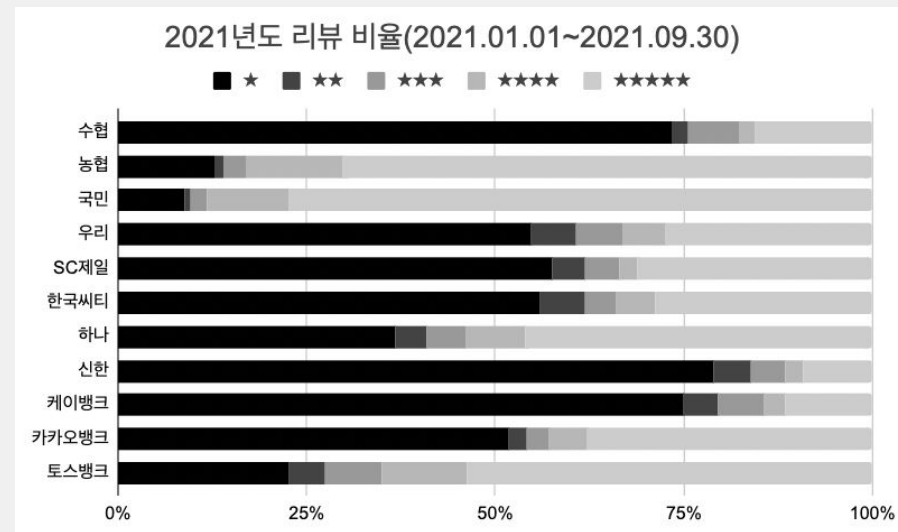
star_4 ▾

star_5 ▾

Csv 파일 -> 스프레드 시트
2021년도 리뷰를 별점 별
시트로 구분

RESULTS

2021년도 리뷰 개수(2021.01.01~2021.09.30)						
은행	★	★★	★★★	★★★★	★★★★★ ★	전체
수협	33	1	3	1	7	45
농협	1,072	101	253	1,090	5,932	8,448
국민	2,257	203	581	2,808	19,882	25,731
우리	1,154	127	129	121	577	2,108
SC제일	366	28	29	15	199	637
한국씨티	413	43	31	38	212	737
하나	790	92	110	169	992	2,153
신한	2,540	158	145	78	296	3,217
케이뱅크	1,999	121	167	74	307	2,668
카카오뱅크	1,481	64	85	143	1,083	2,856
토스뱅크	1,709	354	563	855	4,032	7,513



03

SCORING

씨티은행 리뷰 데이터의
명사/문장 점수 매기기 예시



한국씨티은행의 씨티모바일 앱

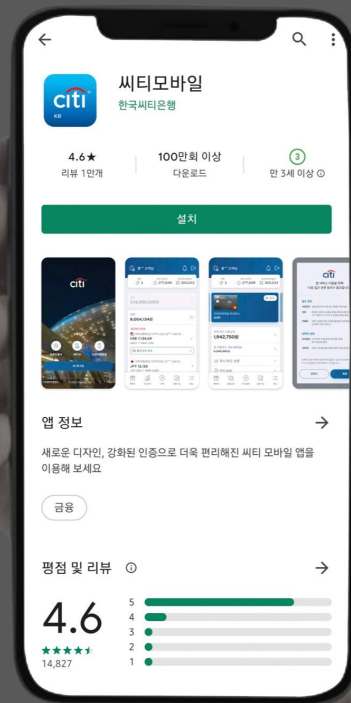


주어진 프로그램을 바탕으로
계산 프로토타입을 만들고,
이를 직접 분석한 것과 비교해
수정 및 보완해 가며 완성도 높임

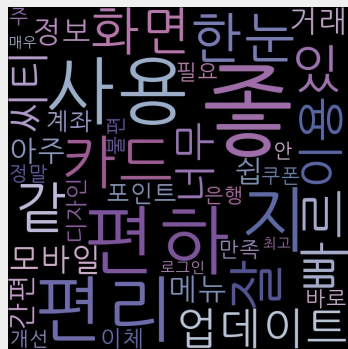


씨티모바일 앱의 2021년 리뷰는 737개, 대상 11개 앱 중 9번째

문장을 직접 보고 비교하는 과정을 포함하므로
지나치게 많은 리뷰 수의 앱을 제외시키고,
비교적 각 별점의 비율이 고른 앱을 선정



각 별점별 리뷰에서 자주 나오는 단어 추출



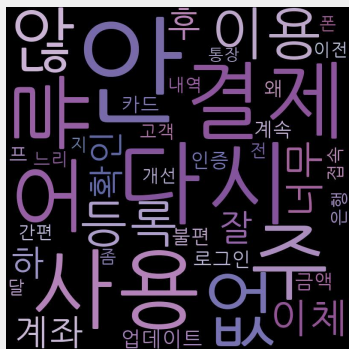
★★★★★



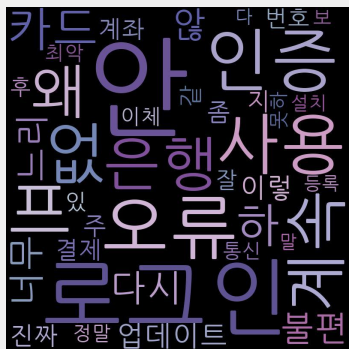
★★★★☆



★★★☆☆



★★☆☆☆



★☆☆☆☆



높은 별점의 리뷰



낮은 별점의 리뷰

경향성 발견 가능.
어휘에 점수를 부여할 때,
어떤 단어가 높은 별점이나 낮은 별점의 리뷰에 자주 등장하는지 확인하고 정성적으로 반영

Q44	A	B	C	D
1	noun	counts		
2	0 앱	237		
3	1 사용	157		
4	2 안	152		
5	3 오류	96		
6	4 카드	78		
7	5 불면	72		
8	6 은행	62		
9	7 인증	62		
10	8 결제	60		
11	9 후	58		
12	10 계좌	54		
13	11 면리	53		
14	12 화면	47		
15	13 이체	45		
16	14 번호	39		
17	15 등록	38		
18	16 때	37		
19	17 이용	35		
20	18 확인	34		
21	19 개선	33		
22	20 설치	30		
23	21 인증서	29		
24	22 로그인	29		
25	23 메뉴	29		
26	24 화각	28		
27	25 통신	27		
28	26 접속	26		
29	27 내역	25		
30	28 조회	24		
31	29 입력	24		
32	30 예러	22		
33	31 속도	22		
34	32 반응	22		
35	33 씨디	22		
36	34 거래	22		
37	35 식제	22		
38	36 전화	21		
39	37 비밀번호	21		
40	38 이후	21		
41	39 한눈	20		
42	40 기능	20		
43	41 정보	20		
44	42 문제	19		
45	43 고객	19		
46	44 필요	18		
47	45 처음	18		
48	46 상태	17		
49	47 금액	17		
50	48 아이디	17		
51	49 비번	16		
52	50 돈	16		

명사 빈도순 추출 - 문제와 해결방안

‘안’ 등 품사가 잘못 나오는 문제

‘안’의 경우 형용사와 동사 앞에서 부사로 인식하도록 한다.

다른 부사가 잘못 나오는 경우 가장 많이 쓰이는 맥락을 확인하고 그에 맞춘다.

부사 혹은 명사가 될 수 있는 단어가 문장의 끝에 나올 때 부사로 인식한다.

단어가 깨져서 끊어져 나오는 문제
(끊어져서 나온 단어 때문에 일부 단어 개수가 지나치게 크게 나옴 ex. 번호, 계좌)

이 분야에서 사용되는 합성어들을 user dictionary에 추가해 준다.

맞춤법 교정기가 교정해주지 못하는 오타,
줄임말 등

‘업뎃,’ ‘업뎃,’ 등 자주 등장하는 경우 user dictionary에 추가해 준다.
이외의 경우에는 맥락 없이 보아도什么的 오타 혹은 줄임말인지 명백한 경우에만 user dic.에 추가해 준다.

Q44	A	B	C	D	E	F	G	H	I	J	K	L	M
1	noun	counts		reinforcer	score				error	etc			
2	0 앱	237		0	0								
3	1 사용	157		0	0								
4	2 안	152						MAG		바로 뒤에 동사형용사/부사 나오는 경우 MAG			
5	3 오류	96		0	-3								
6	4 카드	78		0	0								
7	5 불면	72		0	-3					한눈에 보기 불면 - reinforcer 값 필요할 수 있음			
8	6 은행	62		0	0								
9	7 인증	62		0	0								
10	8 결재	60		0	0								
11	9 후	58		0	0								
12	10 계좌	54		0	0					계좌이체, 계좌번호 등 하나로 되는 것 등록하기			
13	11 면리	53		0	3								
14	12 화면	47		0	0								
15	13 이체	45		0	0								
16	14 번호	39		0	0					계좌번호? 핸드폰번호? 비밀번호? 인증번호?			
17	15 등록	38		0	0								
18	16 때	37		0	0								
19	17 이용	35		0	0								
20	18 확인	34		0	0								
21	19 개선	33		0	1								
22	20 설치	30		0	0								
23	21 인증서	29		0	0								
24	22 로그인	29		0	0								
25	23 메뉴	29		0	0								
26	24 회약	28		0	-5								
27	25 통신	27		0	0								
28	26 접속	26		0	0								
29	27 내역	25		0	0								
30	28 조회	24		0	0								
31	29 입력	24		0	0								
32	30 예러	22		0	-3								
33	31 속도	22		0	0								
34	32 반응	22		0	0								
35	33 씨디	22		0						고유명사, VALUABLE - 별점에 따라 다른 공부정 값 EX. 역시 씨디			
36	34 거래	22		0	0								
37	35 식재	22		0	-4 (변경가능)					식재기 해결방법인 경우 - 별점에 따라 가중치			
38	36 전화	21		0	0								
39	37 비밀번호	21		0	0								
40	38 이후	21		0	0								
41	39 한눈	20		0	2					균형적인 가치, 한눈에 보기 불편한 경우 어떻게 계산할지 고민 필요			
42	40 가능	20		0	0								
43	41 정보	20		0	0								
44	42 문제	19		0	-1								
45	43 고객	19		0	0								
46	44 필요	18		0	-1					나중에 더 보기			
47	45 처음	18		0	0								
48	46 상태	17		0	0								
49	47 금액	17		0	0								
50	48 아이디	17		0	0								
51	49 비번	16		0	0								
52	50 돈	16		0	0								

명사 점수 부여 방식

3명의 합의 하에 정한다

명사는 더하기 점수만 갖는다.

긍정과 부정의 절댓값은 최대 5로 한다.
소수점 한 자리까지 쓸 수 있다.

유사한 개념의 경우 대조하여 맞춘다.
데이터가 쌓이면 상대평가한다.

해당 분류된 단어가 어떤 문장 안에서
사용되었는지, 프로그램을 통해
확인하여 자주 사용되는 맥락을
파악하고 반영한다.

명사 긍·부정 점수 부여 - 문제와 해결방안

일부 명사가 형용사처럼 기능하는 경우

명사 앞에 명사가 나오면 앞의 명사를 형용사처럼 계산하도록 한다.
그 외, 이 분야에서 사용되는 고유한 합성어는 user dic.에 추가해 준다.
Ex. 간편 결제

‘하’ 같은 동사파생접미사, 형용사파생접미사가 붙으면 서술어가 되어
reinforcer 값을 가질 수 있다. Ex. 한눈에 보기 불편하네요: ‘한눈에
보기’라는 긍정적인 명사절을 부정적으로 강화하는 ‘불편(명사) / -
하-(접미사)’

양가적인 단어의 경우

별점에 따라 가중치를 부여한다.

중성이지만 강화될 가치가 있는 경우

더해지기만 할 때는 0, 강화 대상이 될 때는 1을 값으로 갖도록 한다.

지금까지의 나온 것들 중 user dictionary에
넣을 것들 예시

공동인증서, 카카오페이, 아이디, 어플, 앱, 업뎃(업뎃, 앱뎃 등), 먹통,
렉, 간편결제, 계좌번호(계좌 번호), 카뱅, 씨티, 지문인증, 인증번호 등
> 띄어쓰기가 있어도 기능하도록 만들기

문장 긍부정 계산 알고리즘



품사

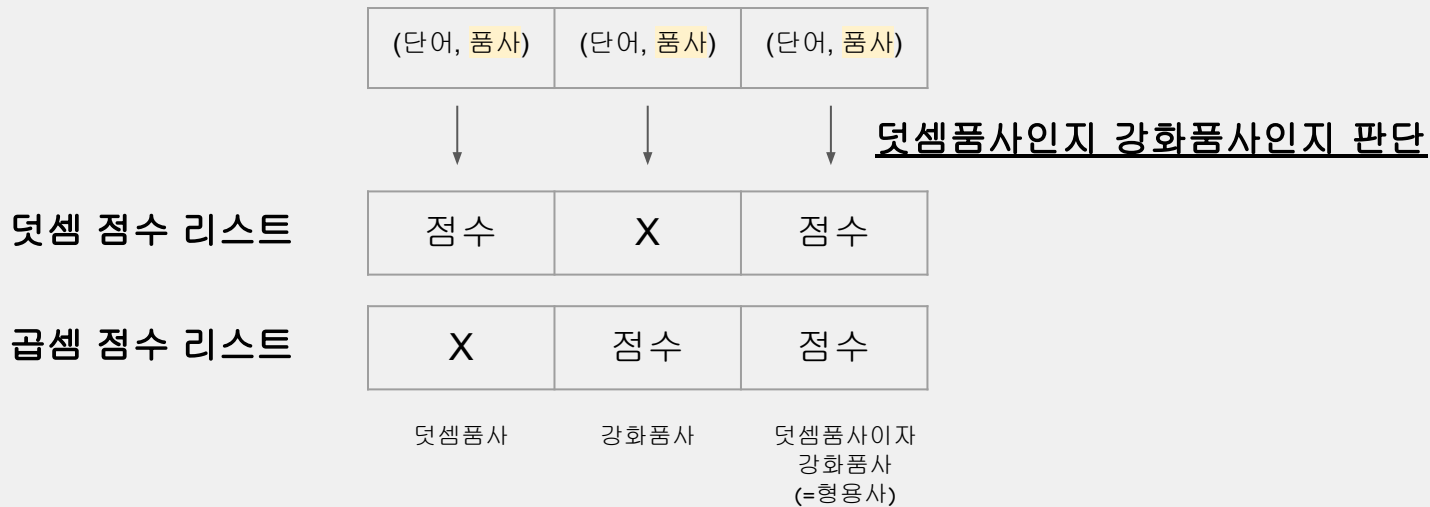
덧셈품사 : NNG(일반명사), NNP(고유명사), VV(동사), VA(형용사), NR(수사), IC(감탄사), 어근(XR),

강화품사 : MAG(부사), VX(보조용언), VA(형용사)

문장 긍정 계산 알고리즘

덧셈품사 : NNG(일반명사), NNP(고유명사), VV(동사), VA(형용사), NR(수사), IC(감탄사), 어근(XR),

VCN(긍정지정사), VCP(부정지정사),
강화품사 : MAG(부사), VX(보조용언), VA(형용사)



문장 긍정정 계산 알고리즘

덧셈 점수 리스트

곱셈 점수 리스트

점수	X	점수
X	점수	점수
덧셈품사	강화품사	덧셈품사이자 강화품사 (=형용사)



덧셈점수와 곱셈점수를 같이
가지는 **형용사에 대한 처리**가
필요!

언제 덧셈점수만 갖고
언제 곱셈점수만 갖는지 판단

문장 긍부정 계산 알고리즘

1 보기 안좋아요

[['보', 'VV'], ['안', 'MAG'], ['좋', 'VA']]

2 한눈에 보기 안좋아요

[['한눈', 'NNG'], ['보', 'VV'], ['안', 'MAG'], ['좋', 'VA']]

보 : 중성 / 안좋 : 부정 / 한눈 : 긍정

1번 문장 : 중성 + 부정 -> 부정

2번 문장 : 긍정 + 부정 -> 중성(?)



형용사가 문장 맨끝에 올 때

형용사는 강화품사가 되고 전체점수에 강화값을 곱한다.

1 쓰기 좋은 앱이다

[['쓰', 'VV'], ['좋', 'VA'], ['앱', 'NNG'], ['이', 'VCP']]

2 쓰기 좋다

[['쓰', 'VV'], ['좋', 'VA']]

1번 문장 : '좋' 이 '앱'을 꾸미는
관형어로 쓰임

2번 문장 : '좋' 이 서술어로 쓰임



형용사 뒤에 명사 성분(일반,고유)이 있으면

형용사는 강화품사가 되고 명사의 덧셈점수에 강화값을 곱한다

형용사 뒤에 명사 성분(일반,고유)이 없으면

형용사는 덧셈품사가 된다

문장 긍정성 계산 알고리즘

형용사를 앞선 규칙에 따라 덧셈품사 or 강화품사 로 설정하면,
점수 리스트 형태는 같은 인덱스에서 덧셈점수나 곱셈점수 둘 중 한 점수만 갖고 있는 형태가 됨

덧셈 점수 리스트

점수	X	점수
----	---	----

곱셈 점수 리스트

X	점수	X
---	----	---

최종 점수

=

\sum 덧셈 점수 리스트

점수	X	점수
X	점수	X

↑
곱셈 점수
리스트

곱셈점수리스트는 덧셈점수 리스트에 들어있는 점수값을 변경하는데 사용됨

곱셈 점수는 맨끝에 오지 않는 이상 무조건 다음 인덱스 덧셈 점수 값에 곱하고,
다음 인덱스 덧셈 점수값이 없다면 다음 인덱스 곱셈 점수 값에 곱해진다. (덧셈 점수값이 없으면 곱셈 점수값이 있다는 뜻)

04

PROJECTS

문제제기와 앞으로의 계획

추후 계획

1. 단문 문장 20개를 임의로 추출하여 설정한 계산 알고리즘으로 금부정 점수를 계산해보고 문장과 결과값을 비교하여 예외상황 확인

→ 예외상황 어떻게 처리할 지?
알고리즘에서 변경/추가할 규칙 생각

2. 장문 문장을 쪼갤 수 있는 방법 고민
(현재는 연결어미, 종결어미로 구분하는 방안을 고안 중)
예시)

원문: 설정 들어가면 아무것도 안되고 먹통이되요

그리고 원하는걸 하기에 너무 난잡하고 찾기도힘들어요

앱자체가 렉도심하고 끊김이 있어요

A: [('설정', 'NNG'), ('들어가', 'VV')]
B: [('아무것', 'NNG'), ('안', 'NNG')], 부정
C: [('먹통', 'NNG'), ('이', 'VCP')], 부정
D: [('원하', 'VV'), ('하', 'VV'), ('너무', 'MAG'), ('난잡', 'XR')], 부정
E: [('찾', 'VV'), ('힘들', 'VA')], 부정
F: [('끊기', 'VV'), ('있', 'VV')], 부정
최종 점수 = A + B + C + D + E + F

3. 덧셈점수도 강화점수도 갖지 않는 단어들은 0점을 부여하여 중성 단어로 보고 1의 방법으로 test 하면서 쓸모없다고 판단되는 명사를 STOP WORDS로 지정

생각해야 할 점

형용사의 점수 계산 방법

-> 품사의 위치로 덧셈/강화품사 정하기에는 문법적 신뢰성이 떨어짐
생각하고 있는 방법 : 부정 형용사면 강화품사, 긍정 형용사면 덧셈품사

강화값이 언제 특정 점수값을 곱할지 전체 점수값을 곱할지

-> 부사의 위치가 자유롭다는 점에서 위치를 기준으로 계산할 시 예외상황 많이 나올 가능성 높음
생각하고 있는 방법 : 어떤 어미가 나오고 뒤에 강화품사가 오면
전체 값을 곱해주거나 일부를 곱하는 식으로 결정하는 방법

단문 리뷰 와 장문 리뷰 결과값의 범위문제

장문 리뷰의 경우 : 여러개의 단문 계산값의 누적값이 최종 점수임
-> 동일한 긍부정 정도를 갖는 단문/복문 리뷰가 점수에서 큰 차이를 보임

END