

어플 리뷰 분석을 통한

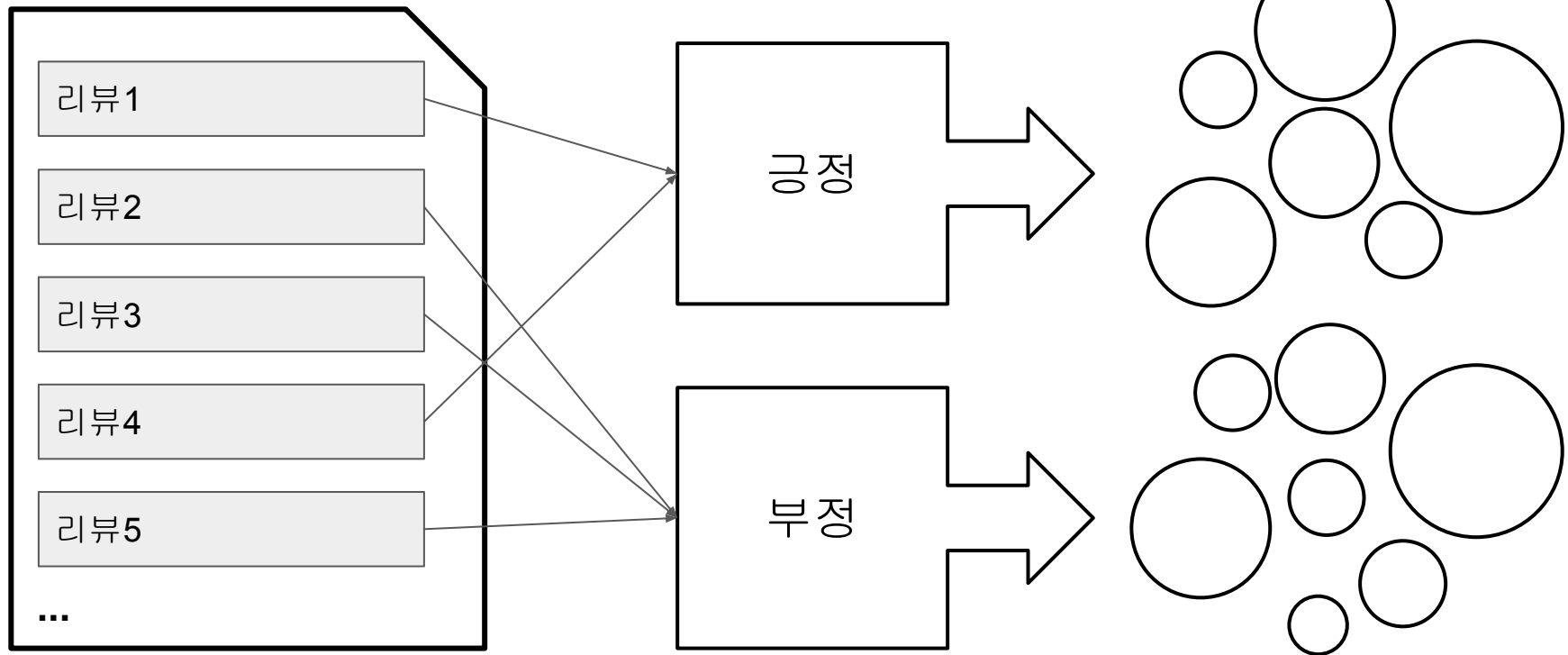
모바일뱅킹어플

개선 방향 제시

4조 이은, 문결, 여윤아

주제 설명

이 작업을 11개 은행에 대해 실시



▶사용자가 좋아하는 요소,
개선되어야 하는 요소 확인

형태소별로 쪼갬



계산할 품사 추출



추출 단어들 리스트에 저장



리뷰1

리뷰2

리뷰3

리뷰4

리뷰5

긍부정 계산값

3.5

-2

6

8

...

리뷰 별점

5점

1점

2점

4점

VS

기존 사전의 점수 수정 및
사용자 정의 추가 구축

계산 규칙 알고리즘 따라
글의 긍/부정 점수 계산

Data 수집

수집 위치 : Google play store

수집 방식 : Google play scrapper 사용

수집 대상 :

- 1) 금융소비자연맹 2021 은행순위 기준, 지역은행 제외 상위 은행들의 모바일뱅킹 앱 리뷰 (11개)
- 2) 최근 6개월 리뷰 (어플 특성상 잦은 업데이트로 오래된 리뷰는 신뢰성이 떨어짐)
- 3) 리뷰별 공감수 (긍/부정 리뷰들 각각에서 도출되는 연관 단어들의 중요성 판단에 기여)
- 4) 리뷰별 별점 (양가적 어휘가 있는 경우 별점에 따라 점수를 선택하도록 하는 방안 고려중)
- 5) 어플의 전체 평점 (리뷰 분석을 통해 도출한 은행 앱의 최종 긍부정점수와 비교)

Data 수집



신한 솔(SOL) - 신한은행 스마트폰뱅킹

파일 수정 보기 삽입 서식 데이터 도구 부가기능 도움말

	A	B	C	D	E
1	userName	score	date	thumbsUpCount	comment
2	황재원	1	2021-09-29	0	사용 5분만에 왜 별점이 3점인지 알겠습니다. 3점도 후해보이네요.
3	미키마우스	1	2021-09-29	0	타은행보다 앱이 보기편한데 계속 튕겨서 이제 이외엔 될 못함
4	캡틴아메리카노	1	2021-09-29	0	앱이 기본적으로 너무 무겁고 일반 고객이 쓰지않는 기능이 너무 많아 불편합니다 솔 라이트로 필요한기능만 넣고 직관적으로 빠르고 가
5	dotobe0929	1	2021-09-29	0	이거 왜 실행도 안되고 튕겨요?
6	Ji Ae Hwang	1	2021-09-29	0	자꾸 튕기고 안눌리고 난리도 아니네요..
7	장종인	1	2021-09-29	0	다른앱을 사용 중에도 광고가 뜨네요—— 신한 솔을 삭제한 후에는 광고가 안뜨구요
8	Coala Cho	1	2021-09-28	0	골드 실버 투자할 때 99%의 확률로 오류가 뜨는데 이거 왜 안고칩니까? 똑같이 시행해서 될 때가 있고 안 될 때가 있다는 것은 명백하게 :
9	방귀할매	1	2021-09-27	1	번호인증이 왜 안 되는거죠? 개인이 만든 어플보다 못하네 진짜
10	동물의 숲	1	2021-09-27	6	다른건 다 모르겠고... 타은행에 비해서 사용하기 너무 불편해요. 통장내역 표시가 입금 출금이 다예요?? 아니 기존 잔고는 도대체 어디서
11	김준제	1	2021-09-27	0	요새 왜이럼? 로그인안되고 튕기고 먼 짓거리를 했길래
12	Demi	1	2021-09-27	0	요즘 왜 이렇게 로딩이 오래걸리고 중간에 네트워크 연결이 끊기는지..
13	I Han	1	2021-09-27	5	비밀번호 입력하면 "지점 상태 오류입니다. 해당 영업점에서 거래 불가입니다. 고객센터로 문의주세요." 도대체 이것은 무슨 오류인가요?
14	kevin lemon	1	2021-09-26	3	홈 화면 변경시 로그아웃되며 변경안됨. 기본 화면 로딩 오래걸리고 멈춤 현상 잦음
15	빈이	1	2021-09-25	0	지문 변경되었다고 몇번을 인증페이지만 보여주나 아 진짜
16	강현규	1	2021-09-25	0	안드로이드 12 베타 나온지가 언제인데 지원할 생각이 없는건지 인력이 딸리는건지 역량이 안되는건지 온갖 금융앱들중에 신한 관련앱들

총 11개 은행에 대한
최근 6개월 구글 플레이스토어 리뷰 크롤링 수행

은행 별 비교

		평균 평점	순위	평점 평가자 수	순위	리뷰 개수	순위	평가자 대비 리뷰 비율	순위
특수은행	수협	3.1	8	308	11	157	11	51.0%	6
	농협	4.0	5	30,656	5	15,753	5	51.4%	5
시중은행	국민	3.9	6	90,509	2	60,086	1	66.4%	2
	우리	3.5	7	18,789	6	7,955	7	42.3%	8
	SC제일	2.6	11	4,203	10	3,119	10	74.2%	1
	한국씨티	4.6	1	14,358	7	8,751	6	60.9%	3
	하나	4.3	2	12,726	8	5,544	8	43.6%	7
	신한	3.0	9	55,176	3	23,230	3	42.1%	9
인터넷 전문은행	케이뱅크	2.8	10	8,657	9	5,076	9	58.6%	4
	카카오뱅크	4.3	2	45,992	4	17,080	4	37.1%	10
	토스뱅크	4.2	4	177,816	1	46,964	2	26.4%	11

토스뱅크) 평가자 수와 리뷰 개수가 최상위권인데 반해 평가자 대비 리뷰 비율이 최하위임.

SC제일은행) 토스와 반대로 평가자 수와 리뷰 개수가 최하위권인데 반해 평가자 대비 리뷰 비율이 최상위임.

토스뱅크) 평점 평가자 수가 타은행 대비 굉장히 많음, 2위인 국민은행과도 약 2배 차이

**추후 은행 어플 별 최종 어플 공부정도 점수 비교 시
은행 어플 별 평가자수와 리뷰 개수 다른 점을 보정해줘야 하지 않을까?**

형태소 분석기 비교

형태소 분석기 비교(komoran, oct, mecab) ☆ 📖 ☁

파일 수정 보기 삽입 서식 테이블 도구 추가기능 도움말 어제 오후 8:25에 마지막으로 수정했습니다

100% W % .0 .00 123 기본값 (Arial) 10 B I S A

<

Komoran이 ‘형태소’ 분석기에 가장 가까움, 모든 형태소 분리 분석. 일부 종결어미를 연결어미로 분석하는 오류 발견됨.

Ok!의 경우 더 큰 덩어리로 나눠줌,
프로그램으로 분석하기 더 용이해
보임. 대상 텍스트로 실험했을 때 몇
가지 오류 발견됨.

Mecab의 경우 Komoran과 유사하지만 어미를 용언에 결합해서 분석해 줌. 연결어미 등을 문장성분 분석에 활용하는 것을 고려하고 있으므로 이 경우 mecab 활용하기 어려울 수 있음.

사전 사용

기존 사전을 활용하되, 검토하고 수정 및 추가할 것

점수 변경 방식 : 3명이 합의하여 점수 확정하는 방식

Sample Analysis

긍정 단문 : “이체가 간편해서 좋아요” (별점4, 공감0)

부정 단문 : “최악의��킹앱 / 시중은행대비 대출 금리 비싼건 덤 / 기타 금융상품도 매력이 없음” (별점1,공감0)

긍정 복문 : “완벽한건 아니지만 다른은행 앱에 비하면 ㄹㅇ 천사다” (별점5,공감4)

부정 복문 : “별하나주기도 아깝습니다” (별점1, 공감0)

형태소의 구분

- 긍부정 평가에 사용하는 형태소: 실질형태소
- 계산 방식 결정에 사용하는 형태소: 형식형태소(문법형태소)

실질형태소: 명사, 대명사, 수사, 관형사, 부사, 동사의 어근, 형용사의 어근, 감탄사

형식형태소: 조사, 접사(접두사, 접미사), 어미(분류 총위 미정, 연결어미/종결어미 활용?)

> 품사 분류 체계가 Komoran과 다르므로 다시 분류할 필요 있음

*형식형태소(+접속사)를 통해 구/절을 묶고 강화 대상을 구분시키는 방안 검토중
(이어진문장, 안은문장 구분 기대)

“이체가 간편해서 좋아요” (별점4,공감0)

이체 / 가 / **간편** / 하 / 아서 / **좋** / 아요
(명사) (주격조사) (어근) (형용사파생접미사) (연결어미) (형용사) (종결 어미)

<평가 대상>

긍정 지정사(VCP), 부정 지정사(VCN)
일반명사(NNG), 동사(VV), 형용사(VA), 부사(MAG),
수사(NR), 감탄사(IC), 어근(XR), 보조 용언(VX)

이체 / 간편 / 좋 = 0 + 긍정 + 긍정 = 긍정

ex) 간편 : +1, 좋 : +3 = 4점 긍정 문장! == 직관적 긍/부정 점수와 비슷
만약 ‘너무 / 좋’ 인 경우 ‘너무’ 강화점수가 ‘좋’만 강화하는 식으로 계산
(리스트 인덱스 활용해서 알고리즘 짤 생각...)

=> 단문은 더하기 점수를 차례대로 더하는 식으로 계산

Komoran 분석 결과

```
===== Computing =====  
이체 --> 더하기:  0.0  
간편 --> 더하기:  1.0  
좋  --> 더하기:  3.0  
좋  --> 곱하기:   1.0  
  
=====
```

<문자 긍정/부정 점수>
score = 4.0, reinforcer = 1.0, final score = 4.0

분석 대상 형태소가 형용사와 명사만 있는 단문의 경우 큰 문제 없음.

다만, ‘너무’등 강화하는 형태소가 들어갈 경우 전체를 크게 강화해 점수 차이가 생김.
> ‘너무’가 ‘좋’만 강화하도록 제한할 필요 있음

“최악의 बैं킹앱 시중은행대비 대출 금리 비싼건 덤 기타 금융상품도 매력이 없음” (별점1,공감0)

[최악의 बैं킹앱] [시중은행대비 대출 금리 비싼건 덤] [기타 금융상품도 매력이 없음]

단문 3개로 나누어서 각각 단문을 앞의 단문 계산법대로 계산하고 전부 더함

[최악의 बैं킹앱] : $-5 + 0 + 0 = -5$ (직관적 점수와 동일)

[시중은행/대비 대출 금리 비싼건 덤] : $0+0+0+0+(-2) + 2 = 0$ (?? 부정문장인데 중립나옴..)
→ ‘(-ㄴ 건) 덤’이라는 명사는 앞문장의 긍부정을 강화하는 역할로 쓰임, 여기서는 0점처리
(긍정으로 쓰이는 경우 구분하는 방법 필요함, 별점을 지표로 활용하는 방안)

[기타 금융상품도 매력이 없음] : $(0+0+2)*(-1) = -2$

Komoran 분석 결과

```
===== Computing =====
최악 --> 더하기:  -5.0
뱅킹 --> 더하기:   0.0
앱 --> 더하기:   0.0
시중은행 --> 더하기:  0.0
대비 --> 더하기:   0.0
대출 --> 더하기:   0.0
금리 --> 더하기:   0.0
비싸 --> 더하기:  -2.0
비싸 --> 곱하기:   1.0
덤 --> 더하기:   0.0
기타 --> 더하기:   0.0
금융 --> 더하기:   0.0
상품 --> 더하기:   0.0
매력 --> 더하기:   2.0
없 --> 더하기:   0.0
없 --> 곱하기:  -1.0
=====





<문자 금/부정 점수>
score = -5.0, reinforcer = -1.0, final score = 5.0
```

코모란처럼

모두 더하고 강화자를 한 번에 곱하는 경우
왼쪽과 같은 문제 생김
(부정을 추가하는 것이 아니라 긍정으로 만듦)

강화어가 강화하는 대상 제한시킬 필요 있음

“완벽한건 아니지만 다른은행 앱에 비하면 ㄹㅇ 천사다.”
(별점5,공감4)

완벽 / 한 / 것 /  은 / 아니  / 지만 / 다른 / 은행 /
앱 /  에 / 비하  / 면 / ㄹㅇ / 천사 / 다

*'-은 아니 / -에 비하'의 경우 여타 부정지정사와 달리 한정된 것을 부정하므로 -1이 아닌 0을 곱하도록?
특별한 어구는 따로 사전에 넣어서 점수 부여, 지정된 계산방식으로 계산

$$((+5) * 0 + 0 + 0 + 0) * 0.5 + 2 * 2 = 4 \text{ (직관적 점수와 비슷)}$$

*'아니다'가 부정하는 대상 한정시켜서 묶는 방법 구해야 함

*형용사 곱하는가, 더하는가?

Komoran 분석 결과

```
===== Computing =====  
완벽 --> 더하기:  5.0  
아니  --> 곱하기: -1.0  
은행  --> 더하기:  0.0  
앱    --> 더하기:  0.0  
르ㅇ   --> 곱하기:  1.0  
천사   --> 더하기:  2.0  
  
=====
```

<문자 금/부정 점수>
score = 7.0, reinforcer = -1.0, final score = -7.0

보조사의 의미 제한이 고려되지 않고, 강화자의 강화 대상이 제한되지 않아 전체 문장이 부정으로 평가됨

“별하나주기도 아깝습니다” (별점1, 공감0)

별하나 주기도 아깝습니다



‘별하나, 일점, 만점, 빵점...’ 등 별점 또는 숫자에 대한 단어들에 대해 긍/부정 점수 부여

별하나 / 주 / 기 / 도 / 아깝 / 습니다
(명사) (본용언) (명사형전성어미) (보조사) (형용사) (종결어미)



부족: ‘아깝’ 처럼 부정의 의미 갖는다고 생각하지만

별다섯개가 부족하다 -> 극찬 / 어떤 기능이 부족하다 -> 부정 (??)

특정 표현이 가리키는 대상에 대하여 다른 점수 계산 방식이 필요하지 않을까??

“별하나주기도 아깝습니다” (별점1, 공감0)

별하나 / 주 / 기 / 도 / 아깝 / 습니다
(명사) (본용언) (명사형전성어미) (보조사) (형용사) (종결어미)

ex)

별 다섯 / 부족 / 평점5 (별 다섯개 주기^엔 부족합니다 vs 별 다섯개 줘^도 부족합니다)
평점 5인데 별다섯 부족 -> 극찬 / 평점 5점 아는데 별다섯 부족 -> 3~4점정도의 긍정(5점만점)

=> 별점을 이런 경우에 판단 근거로 사용하면 좋을 듯

$$-5 + (-1) = -6 \text{ (직관적 점수와 비슷)}$$

Komoran 분석 결과

```
===== Computing =====  
별 --> 더하기:  0.0  
주 --> 곱하기:  1.0  
아깽 --> 더하기:  3.0  
아깽 --> 곱하기: -1.0  
  
=====
```

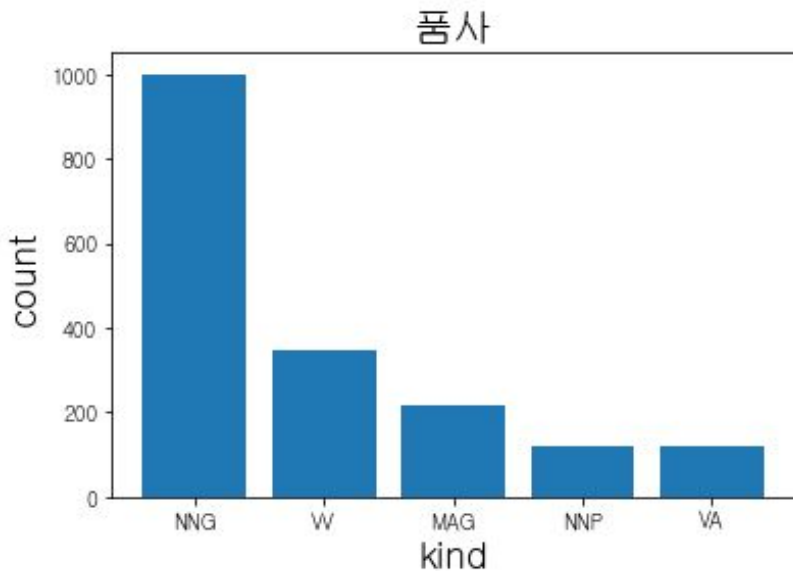
<문자 긍정/부정 점수>
score = 3.0, reinforcer = -1.0, final score = -3.0

수사가 평가에 포함되지 않음, 서술어(강화어)가 서술하는 대상 제한되지 않음.

특수한 표현이 분석되지 않음(user dictionary로 해결 가능한 부분)

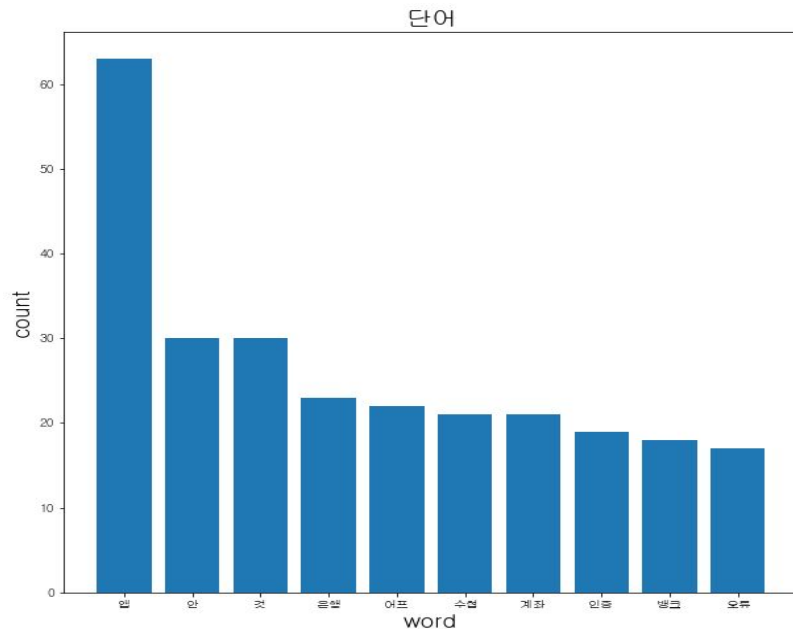
단, ‘별 하나 주-’는 행위가 부정적으로 평가되더라도 부정 강화자가 이를 긍정으로 만들 우려가 있음.
따라서 user dic.을 활용할 경우 ‘별하나 주기도 아깽’ 전체를 지정해 줘야 함.

수협은행 어플 리뷰에 대한 품사/단어 빈도



품사별 빈도 수

-> 일반명사 > 동사 > 일반부사 > 고유명사 > 형용사



가장 많은 빈도수를 갖는

명사 단어들에 대한 빈도수

10개 이상 나타나는 단어들만 그래프에 표시

일반/고유/의존 명사 포함

(앱>안>것>은행>어플>수협>계좌>인증>뱅크>오류)

명사(일반명사 + 고유명사)에 대한 워드클라우드 (수협은행 리뷰)



- ‘어프’ 등은 사전에 등록되지 않은 단어가 분리된 것으로 여겨짐. 워드클라우드를 활용하면 user dic.의 표제어를 등록하는데 유용하게 활용할 수 있을 것으로 예상됨
- 명사만 활용했을 때 워드클라우드로서 의미 있는 결과를 얻을 수 있었음
- 기능에 관한 단어와 부정적으로 볼 수 있는 단어가 특히 눈에 띄음
(ex. 인증, 업데이트, 로그인 / 최악, 오류)
- ‘필요’ 등의 단어 앞에 등장하는 빈출 단어들을 모아 분석할 수 있을 경우 개선 방안 제안에 유용할 듯
- ‘안’의 경우 명사가 아닌 것이 명사로 인식됨. 어떤 맥락에서 어떤 품사를 우선적으로 인식시킬지 지정해줄 필요 있음

- 생각할 것 메모

- 이어진문장 / 안은문장 각자 말뭉치 묶는 방법 필요 > 구분은 어떻게 할 건지(연결어미마다 할당?)
- 프로그램에서 문장성분을 분석해 서술어/부사어가 설명하거나 강화시키는 대상을 한정시킬 수 있을지 (인덱스?)
- 보조사(-은 등)의 역할을 어떻게 표현할 것인지
- '아니다'가 앞의 내용을 무효화시키는지, 부정하는지 (0을 곱한다 or -1을 곱한다) ('완벽한 것은 아니다' 부정정도?)
- (문장성분 사용하는 경우) 국어에서 명사가 명사 앞에서 형용사처럼 사용될 때를 어떻게 구분할지
- (연결어미로 복문을 구분할 경우) '-다'를 코모란에서 연결어미로 판단했는데, 명백히 종결어미인 경우를 어떻게 구분시킬지 *연결어미로 구분할 경우 oct 사용 어려울듯.
- 별하나도 아깝, 별다섯개도 부족 > 긍정과 부정 평가 문제가 됨. 일단 해당 표현은 자주 등장하므로 표제어 등록.
(구분시키는 방법: 별점 활용. 별점 5점인데 별다섯/부족 등장하면 긍정, 아니면 애매함)
- 중립 기준을 0이 아니라 1로 삼는 것은? (중립적 가치이지만 곱했을때 0이면 안 되는 것. ex) 시간 아까움)
- 긍정과 부정 갈리는 경우 별점을 고려하여 긍부정 점수 선택적으로 사용되도록 알고리즘
- 다른 좋은 은행 쓰면 되지 왜 이거 씀? 처럼 직접적으로 부정어가 등장하지는 않는 경우 어떻게 할까?
'다른 좋은'의 문제일까? > 써보니 모임통장 말고 다른 좋은 기능도 많네요~ 의 경우에는 어떻게 할까?
- 구절 분리 패키지: 부정 강화 하나가 문단 전체에 영향을 미치는 위험 분산 가능
- '모임통장'처럼 개별 은행이 제공하는 특이 기능 등: 워드카운팅으로 자주 등장하는 경우 사용자 사전에 정의
- 긍정적인/부정적인 리뷰에서 자주 나오는 어휘 취합하면 평가에 도움이 됨