

맥주 리뷰 데이터를 활용한 맥주에 대한 정보 제공 및 시각화

IT경영학과 2018314004 김남훈

IT경영학과 2018314015 서민혁

IT경영학과 2018316031 정찬호

IT경영학과 2018314010 남도형

CONTENTS



01

연구의 필요성



02

연구 목적



03

연구 방법 및 연구 결과



04

향후 방향성 제시

01 연구의 필요성



" 항상 먹던 맥주나 마실까..? "

" 맥주가 너무 많아서 고르지 못하겠네 "

" 이 맥주는 무슨 맛이지? "



맥주를 구매 할 때,
외관만으로는
그 맥주에
대한 정보를 알기 힘들

맥주에 대한 정보를
얻을 수 있는 SNS,
커뮤니티 공간이 적음



맥주 선택의 어려움 해결 필요

01 연구의 필요성



와인 정보 및 소비자 리뷰를 제공하는 온라인 와인 커뮤니티 앱

- 1. 와인 검색 및 정보** - 사용자가 원하는 와인을 검색하여 해당 와인의 품종, 원산지, 생산자, 가격 등의 정보와 함께 커뮤니티의 회원들이 작성한 리뷰와 평점도 확인할 수 있다.
- 2. 와인 평가 및 리뷰** - 자신이 시음한 와인에 대한 평가와 리뷰를 작성할 수 있다. 이를 통해 다른 사용자들은 와인의 맛, 향, 특징 등에 대한 정보를 얻을 수 있다.
- 3. 스캔 기능** - 와인 레이블을 스캔하여 해당 와인에 대한 정보를 바로 확인할 수 있는 기능을 제공한다. 와인 레이블을 스캔 후 사용자에게 관련 정보와 리뷰를 제공한다.

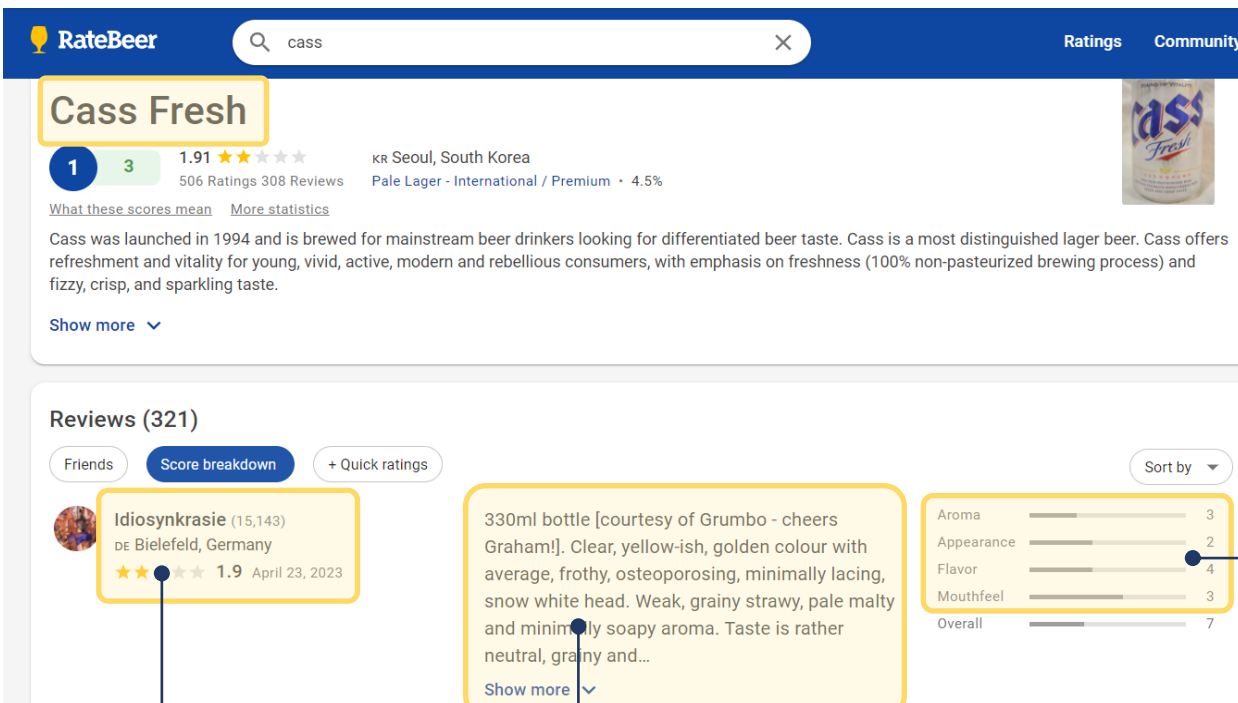
02 연구 목적

각 맥주의 특징들을 **키워드**로
제공함으로써 소비자들에게
편리성 증대

맥주에 대한 사용자
리뷰 데이터를 활용하여
소비가 선택한 맥주와
비슷한 맥주 제공

본 연구는 각 맥주의 특징을 한 눈에 알아보기 쉽게 정보를 제공함으로써,
소비자의 맥주 선택의 폭을 넓히고자 한다.

RateBeer - 해외 맥주 리뷰 사이트



The screenshot shows the RateBeer website interface. At the top, there's a search bar with 'cass' entered. Below it, the 'Cass Fresh' beer profile is displayed, including its rating (1.91 stars), location (KR Seoul, South Korea), and a brief description. A review by 'Idiosynkrasie' is highlighted, showing the reviewer's profile, the review text, and a detailed score breakdown for Aroma, Appearance, Flavor, Mouthfeel, and Overall.

고객 정보,
평점, 날짜

상세 리뷰

크롤링할 데이터

편의점 맥주 중 리뷰가 10개 이상인 맥주 69개

1. 맥주명
2. 고객정보(고객명, 국적, 총 리뷰수, 전체 평점)
3. 상세 리뷰
4. 맥주 평가 지표 4요소

맥주 평가 지표	평가 지표 뜻
Aroma(0~10)	맥주의 향
Appearance(0~5)	맥주를 따랐을 때 외관(거품, 색)
Flavor(0~10)	맥주의 맛
Mouthfeel(0~5)	목넘김

03 연구 방법

크롤링 한 데이터

	ID	rating	date	beer_name	review	Aroma	Appearance	Flavor	Mouthfeel
0	vinivini(9,584)	2.9	December 29, 2022	Kloud Original Gravity	Straight from the can. Clean, malty, bready & ...	5	4	6	4
1	italianjohn(5,521)	2.8	September 6, 2021	Kloud Original Gravity	Clear golden blonde pour with a thin white hea...	5	3	5	3
2	orda(296)	3.9	August 29, 2021	Kloud Original Gravity	Appearance: golden amber; slight head Taste: d...	10	5	10	5
3	Radek Kliber(8,125)	2.4	June 5, 2021	Kloud Original Gravity	Bottle 330 ml Aroma: nice, fresh, on dry sid...	6	3	4	2
4	geologyguy(2,602)	2.3	May 26, 2021	Kloud Original Gravity	Smell is light of malt...taste is a that malt ...	4	3	4	3
...
68571	beerman991979(136)	3.2	September 2, 2000	Warsteiner Premium Verum	I didn't think this was that bad. Yeah it doe...	6	3	6	3
68572	turtlehead(111)	2.9	August 28, 2000	Warsteiner Premium Verum	A yeasty, almost strong pilsner with a pleasan...	8	3	6	2
68573	Aubrey(3,516)	2.9	August 24, 2000	Warsteiner Premium Verum	Initially sharp, but then fizzy and bitter. A ...	5	3	5	3
68574	Sammys(427)	2.9	July 20, 2000	Warsteiner Premium Verum	This was good for a Pilsner, I suppose, but la...	7	3	6	3
68575	PhillyBeer2112(3,413)	3.8	May 31, 2000	Warsteiner Premium Verum	One of the best pilsners - and very identifiab...	7	5	7	4

68576 rows × 9 columns

03 연구 방법

데이터 전처리 진행

1. 소문자화

Straight from the can. → straight from the can.

2. 특수문자 제거

bready & grassy. → bready grassy

3. 토큰화

straight from the can → ['straight',
'from',
'the',
'can',
...]

4. 의미없는 불용어 제거

Smell is light → smell light

5. PorterStemmer를 이용하여 어간추출

원본 단어: really, 어간 추출 결과: realli
원본 단어: enjoyed, 어간 추출 결과: enjoy

6. 2가지 척도 5점만점으로 통일

df['Flavor'] / 10 * 5

03 연구 방법

전처리 후 데이터

Unnamed: 0		ID	rating	date	beer_name	review	Aroma	Appearance	Flavor	Mouthfeel
0	0	vinivini(9,584)	2.9	December 29, 2022	Kloud Original Gravity	straight clean malti breadi grassi light sweet...	2	4	3	4
1	1	italianjohn(5,521)	2.8	September 6, 2021	Kloud Original Gravity	clear golden blond pour thin white head spot l...	2	3	2	3
2	2	orda(296)	3.9	August 29, 2021	Kloud Original Gravity	appear golden amber slight head tast decent li...	5	5	5	5
3	3	Radek Kliber(8,125)	2.4	June 5, 2021	Kloud Original Gravity	bottl aroma nice fresh dri side feel macro vib...	3	3	2	2
4	4	geologyguy(2,602)	2.3	May 26, 2021	Kloud Original Gravity	smell light malt tast malt kinda tast like lig...	2	3	2	3
...
68571	68571	beerman991979(136)	3.2	September 2, 2000	Warsteiner Premium Verum	think bad yeah fit pilsner style hell good bee...	3	3	3	3
68572	68572	turtlehead(111)	2.9	August 28, 2000	Warsteiner Premium Verum	yeasti almost strong pilsner pleasant aroma mu...	4	3	3	2
68573	68573	Aubrey(3,516)	2.9	August 24, 2000	Warsteiner Premium Verum	initi sharp fizzi bitter littl rough palat alm...	2	3	2	3
68574	68574	Sammys(427)	2.9	July 20, 2000	Warsteiner Premium Verum	good pilsner suppos lack bodi opinion refresh ...	3	3	3	3
68575	68575	PhillyBeer2112(3,413)	3.8	May 31, 2000	Warsteiner Premium Verum	one best pilsner identifi bitter alway pleasan...	3	5	3	4

68576 rows x 10 columns

03 연구 방법

전처리 후 워드 클라우드



03 연구 방법

벡터화 (Text Vectorization)

- 전처리 된 리뷰 데이터를 벡터화

CounterVectorizer

```
array([[0, 1, 1, 1, 0, 0, 0, 1, 0, 1],  
       [0, 1, 0, 1, 0, 0, 2, 1, 0, 1],  
       [1, 0, 0, 0, 0, 1, 0, 1, 1, 0],  
       [0, 1, 1, 1, 0, 0, 0, 1, 0, 1],  
       [0, 1, 0, 0, 1, 0, 0, 1, 0, 0]])
```

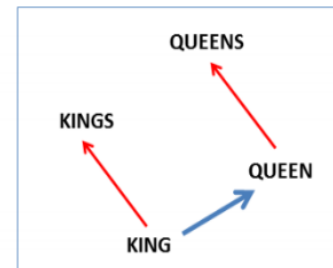
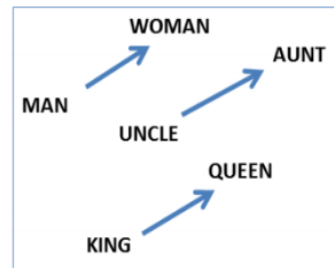
문서 집합에서 단어 토큰 생성
각 문서의 **토큰 출현 빈도**로
bow 인코딩 벡터 생성

TF-IDF

	I	love	dogs	hate	and	knitting	is	my	hobby	passion
Doc1	1	1	1							
Doc2	1		1	1	1	1				
Doc3					1	1	1	2	1	1

단어의 가중치를 조정한 bow
인코딩 벡터 생성
※ 문서에 공통적으로 들어있는
단어는 문서 구별 능력이 떨어진다고
보아 **가중치를 축소**

Word Embedding

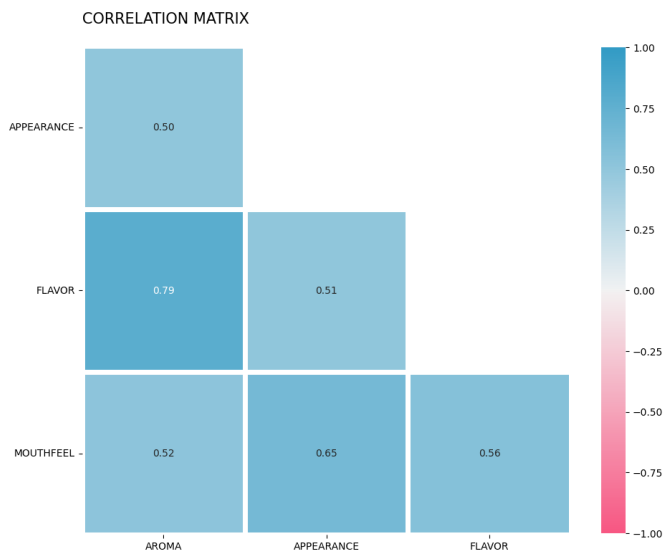


각 단어를 고차원 공간에서 밀집된
벡터로 할당함으로써 단어의
의미와 관계를 표현

03 연구 방법

상관관계 분석 및 다중공선성 확인

- 4가지 지표에 대한 Correlation Matrix 시각화 및 다중공선성 결정계수 출력
- 모든 지표가 상관관계가 있고, 특히 Flavor(맛)과 Aroma(향)의 상관관계가 높음
- VIF Factor 값들이 모두 10 이상으로, 다중공선성이 발생
- 리뷰 자체가 4가지 지표로 나눈 것이 아닌 전체 리뷰이기에 전체 평점을 사용하는 것이 좋을것으로 판단



[Correlation Matrix]

	VIF Factor	features
0	14.812879	Aroma
1	15.596450	Appearance
2	17.566905	Flavor
3	16.855467	Mouthfeel

[지표 별 다중공선성 결정계수]

03 연구 방법

평점 라벨링 및 로지스틱 회귀분석

- 평점 중 1,2점은 부정, 3,4,5점은 긍정 리뷰로 이분화
- 전체 리뷰에 대한 Tf-Idf를 통해 벡터화
- 벡터화 된 리뷰와 라벨링 된 평점을 이용하여 로지스틱 회귀분석 진행 -> 맥주별 회귀계수 추출

Sapporo Premium Beer		
	top_words	회귀계수
0	smooth	2.202364
1	caramel	1.317287
2	citrus	1.130529
3	strong	1.120276
4	refresh	1.116946

Kloud Original Gravity		
	top_words	회귀계수
0	crisp	0.530809
1	smooth	0.499721
2	lite	0.496238
3	bavarian	0.458367
4	cuisine	0.396737

[맥주 관련 단어 별 회귀계수]

03 연구 결과

코사인 유사도 및 추천 결과

- 회귀계수를 활용한 코사인 유사도 계산 및 출력


beer_type	Kloud Original Gravity	Asahi Super Dry	Tsingtao	Heineken	Kirin Ichiban	sapporo Premium Beer / Draft Beer	Stella Artois	Guinness Draught	Kronenbourg 1664 Blanc	Pilsner Urquell	...	Beck's	Hoegaarden Rosée
beer_type													
Kloud Original Gravity	1.000000	0.122228	0.120459	0.108313	0.125984	0.102792	0.094567	0.059279	0.093341	0.091558	...	0.087084	0.028131
Asahi Super Dry	0.122228	1.000000	0.388719	0.343490	0.338171	0.321873	0.335624	0.242308	0.218461	0.278099	...	0.300506	0.128859
Tsingtao	0.120459	0.388719	1.000000	0.364404	0.306570	0.290868	0.349718	0.254658	0.274519	0.271278	...	0.319572	0.132531
Heineken	0.108313	0.343490	0.364404	1.000000	0.325926	0.290310	0.420015	0.333382	0.263045	0.340576	...	0.349459	0.115530
Kirin Ichiban	0.125984	0.338171	0.306570	0.325926	1.000000	0.281081	0.325208	0.208742	0.235768	0.224335	...	0.282330	0.111497
Sapporo Premium Beer / Draft Beer	0.102792	0.321873	0.290868	0.290310	0.281081	1.000000	0.304521	0.228019	0.188840	0.187885	...	0.260254	0.115384
Stella Artois	0.094567	0.335624	0.349718	0.420015	0.325208	0.304521	1.000000	0.311218	0.261146	0.339782	...	0.357030	0.133318
Guinness Draught	0.059279	0.242308	0.254658	0.333382	0.208742	0.228019	0.311218	1.000000	0.227512	0.307169	...	0.265861	0.095354
Kronenbourg 1664 Blanc	0.093341	0.218461	0.274519	0.263045	0.235768	0.188840	0.261146	0.227512	1.000000	0.191488	...	0.225223	0.195520
Pilsner Urquell	0.091558	0.278099	0.271278	0.340576	0.224335	0.187885	0.339782	0.307169	0.191488	1.000000	...	0.290244	0.095505

[코사인 유사도 출력값]

맥주 이름을 입력하세요: Asahi Super Dry
유사한맥주1. Tsingtao
유사한맥주2. Stella Artois
유사한맥주3. Carlsberg Pilsner

[유사 맥주 추천 결과]

[맥주이름]

 Asahi Super Dry

[정보]

★ : 3.5/5

알콜 농도 : 5%



Positive



Negative

[비슷한 추천 맥주]

Guinness | Terra | Heineken

[최종 출력 화면]

04 향후 연구 방향성 제시

OpenCV를 활용한 맥주 객체 탐지

- 객체 탐지를 위해 많이 사용되는 OpenCV를 활용하여, 카메라를 맥주에 비췄을 때, 그 맥주가 어떤 맥주인지 구별하는 탐지 기능 구현 가능
- 해당 기능으로 맥주를 구별한 후, 그 맥주에 대한 기본적인 정보 및 본 연구에서의 분석 결과들을 보여줄 수 있음



[맥주 객체탐지 예시 출력 화면]

04 향후 연구 방향성 제시

본 기능을 활용한 앱 개발

- 본 연구의 결과물과 같이, 소비자들이 손쉽게 맥주에 대한 정보를 얻을 수 있는 어플이나 커뮤니티 등 매체의 부재가 있음
- 이를 시장성이 있다 판단하고, 해당 기능을 구현하는 앱으로 개발하여 소비자의 니즈를 충족시킬 수 있을 것으로 보임



05 참고 문헌

앱 마켓 리뷰 데이터 분석을 통한 배달 앱의 사용자 만족도 연구: 배달의민족과 쿠팡이츠를 중심으로

: <https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11125662>

텍스트 전처리

: <https://wikidocs.net/21707>

자연어처리, Word Embedding

: <https://jh2021.tistory.com/5>

TensorFlow - 텍스트 전처리 학습

: https://www.tensorflow.org/text/guide/word_embeddings?hl=ko