

선형대수학



뉴스 댓글 기반 이모티콘 생성

윤지영 이지윤 이수린 임지훈 채희지

INDEX



주제 선정 및 배경

감정분류

욕설탐지

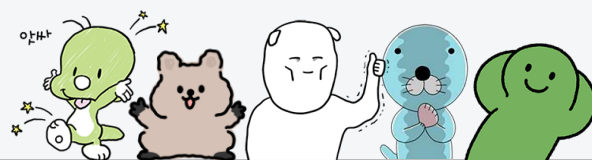
LDA

GAN

DCGAN구현을 통한
이모티콘 생성

분석

구현



1. 주제 선정 및 배경

01. 주제 선정 배경



주제 선정



지난 4월 네이버는 뉴스 서비스에서 **감정 이모티콘을 모두 없애고** 추천 스티커로 대체
네티즌의 불만을 삼



주제 선정



(이재진 한양대 미디어커뮤니케이션 학과 교수 의견 중)

기사에 대한 의사 표시를 하는 방식으로 감정 버튼을 애용해온 이들이
소통 창구를 빼앗긴 느낌이 드는 것은 당연.



네이버 측에서 오히려 **다양한 층위의 감정**을 좀 더 적극적으로 표출할 수 있도록
대안을 만드는 게 좋을 것!

01. 주제 선정 배경



주제 선정



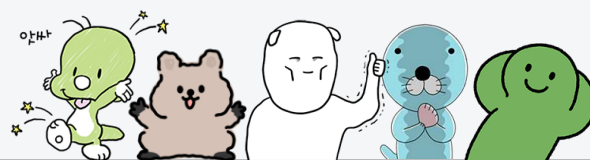
네이버 기사의 특징 및 현황

감정분포

기사의 내용마다
전혀 **다른 감정표현**의 우세

감정표현 불균형

다양성을 반영하기에
현저히 적은 감정표현



2. 감정분류

02. 감정분류



댓글 시각화

감정 대분류에 따른 7가지 감정

기쁨, 슬픔, 놀람, 분노,
혐오, 공포, 중립

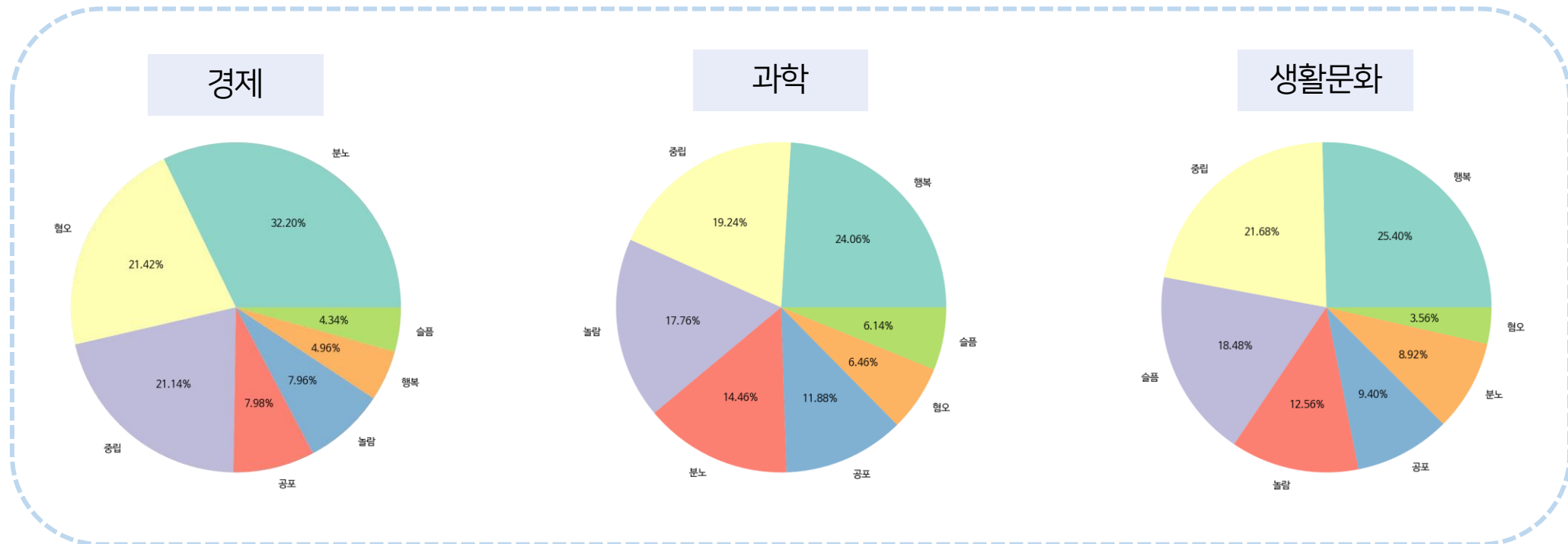


각 분야의 댓글 별 감정 상태를 **대분류**에 따라 나누고
시각화 진행

02. 감정분류



분야별 감정분류 시각화

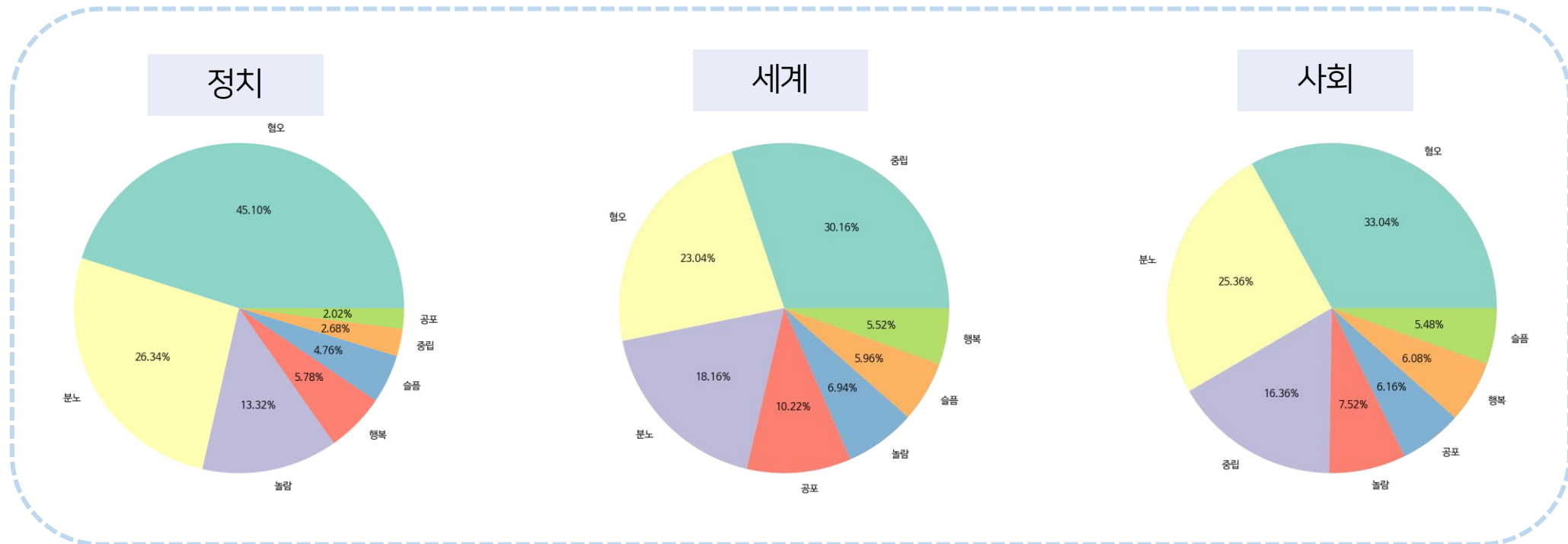


과학과 생활문화에서는 부정적인 감정과 긍정적인 감정이 **비교적 고르게** 분포 되어 있음

02. 감정분류

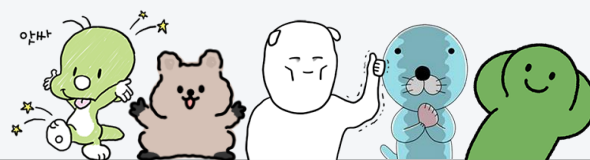


분야별 감정분류 시각화



혐오와 분노가 모든 감정의 절반 이상을 차지함

감정의 불균형이 심함



3. 욕설탐지

문제점 발견



욕설이 담긴 댓글의 감정을 **정확히 분류하지 못하는 문제점** 발견!

데이터셋 자체의 문제

혐오를 드러내는 내용임에도
'ㅋㅋㅋ'가 많아
'행복'으로 **잘못 라벨링된** 경우 多

데이터셋에 없는
새로운 혐오표현을 인식하지 못함

혐오와 중의적 표현이 함께 사용된 경우

좌빨놈 → 혐오
좌빨놈들 덕에 나라 잘 돌아간다 → 중립

미친것 → 혐오
미친것이 날뛰니 Y 지지율 올라가겠다^^
→ 행복

욕설의 형태를 변형하여 쓴 경우

빨개이들 특징 손절 잘하는거 Y도
그렇고 K도 그렇고 조만간 빵 ㄱ ㄱ ㄱ
→ 중립

입 빵긋만해도 정치권에 조우우웃 되는
정치인들썩
→ 중립



문제점 발견



특정 대상을 비난하는 의도의 비속어 표현을 사용했다는 점에서
혐오(부정적) 감정에 더욱 가까울 것이라 판단



추가적으로 욕설탐지를 진행하여
오분류된 부정적 댓글을 찾기로!



전처리



4. FastText 모델을 활용한 임베딩

학습 방식 : Skip-Gram (**중심 단어**로부터 주변 단어를 예측)

N-Gram : 3~6

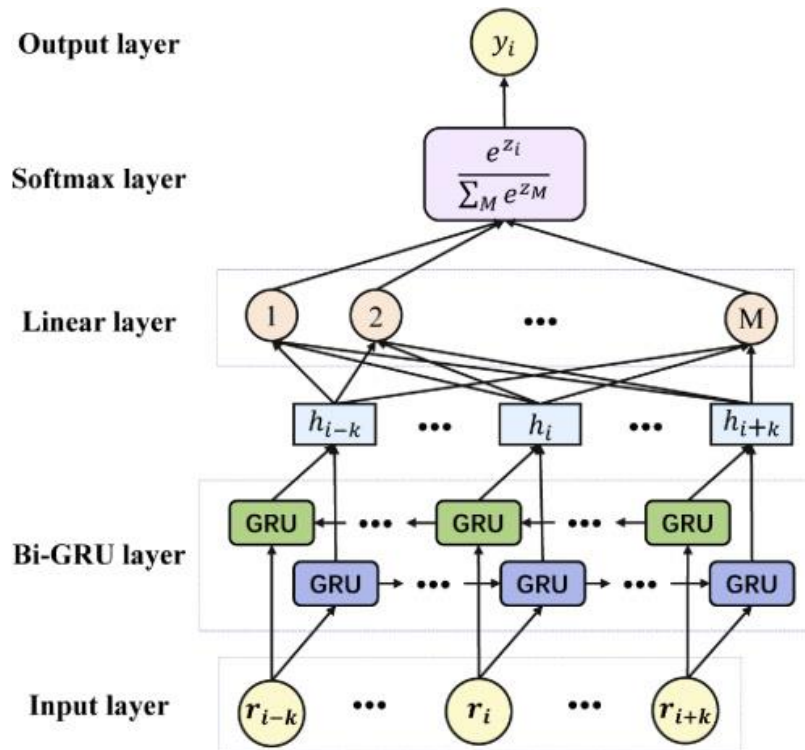
| 단어 | N | BoW |
|-------|---|----------------------------|
| 선형대수학 | 3 | <선, 대, 수, ..., _학, 학, > |
| | 4 | <선, 대, 수, 학, ..., _학, 학, > |
| ⋮ | | |
| 선형대수학 | 6 | <선, 대, 수, 학, ..., _학, 학, > |

Skip-Gram에 대한 자세한 내용은

딥러닝팀3주차 클린업 참고!



모델학습



Bidirectional GRU

온라인 텍스트의 **특정 단어를 탐지**하는 task의 경우,
문자 단위 임베딩+bi-lstm/gru
구성이 고성능을 낸 사례들이 많다는 연구결과 참고

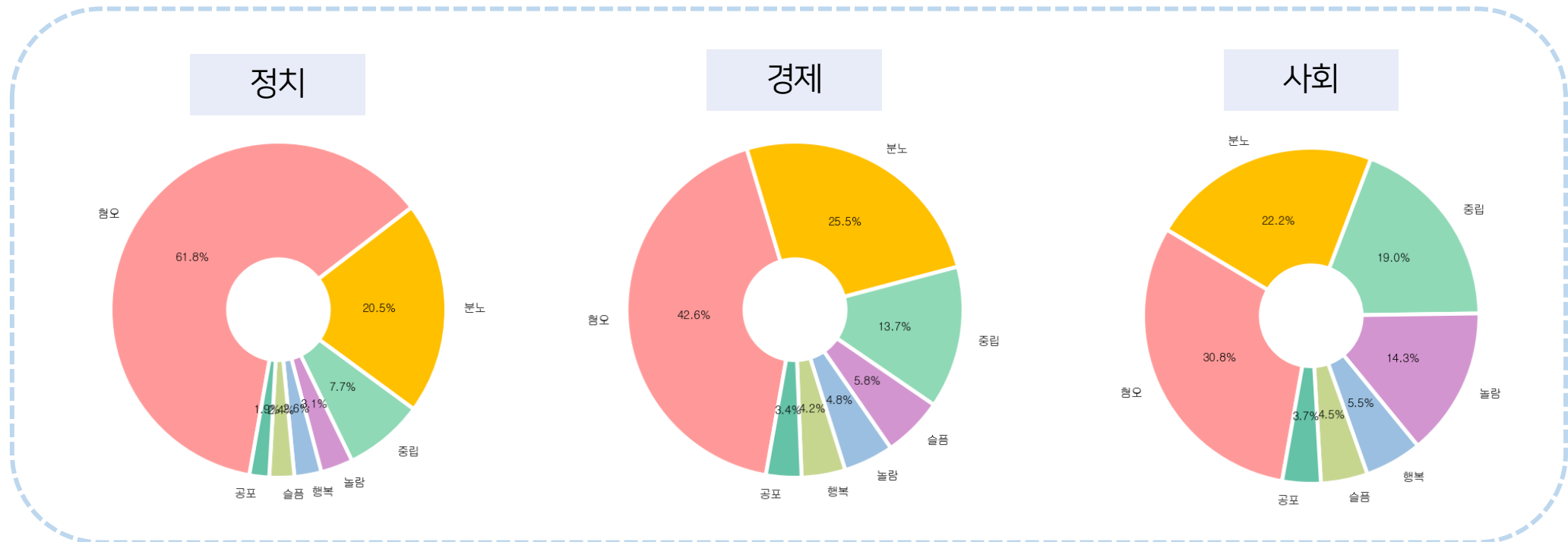
GRU의 자세한 내용은
딤러닝팀3주차 클린업 참고!



03. 욕설탐지



결과

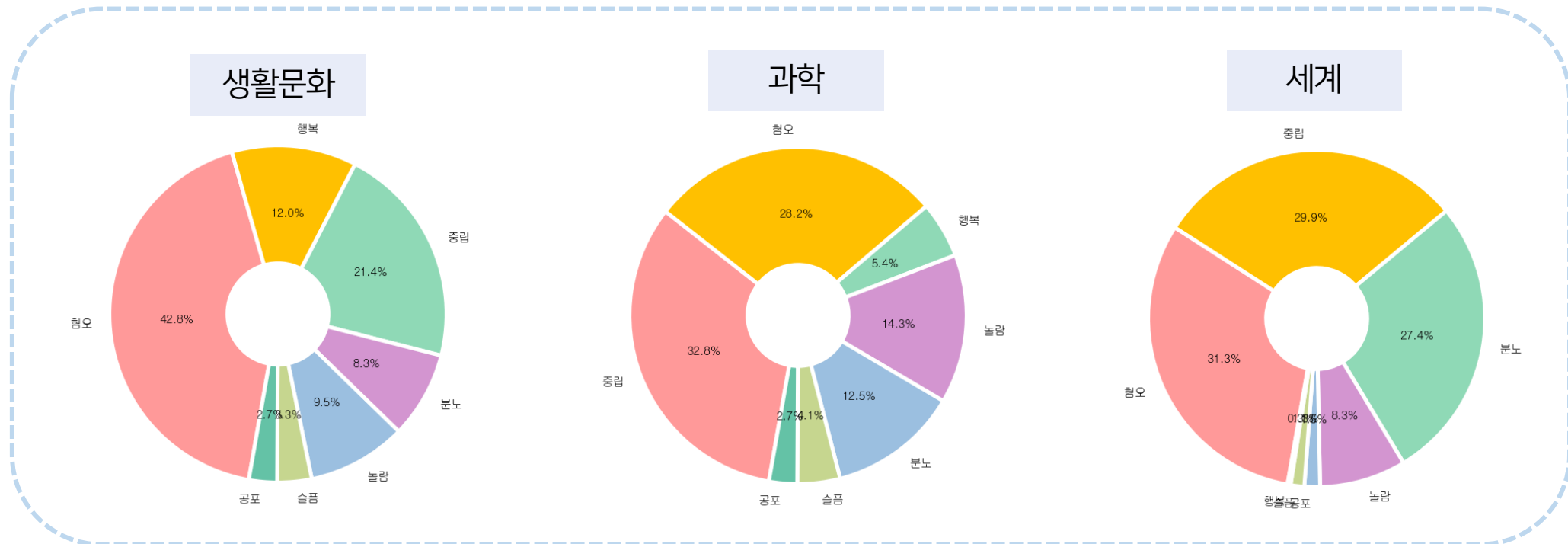


혐오와 분노가 많은 비율을 차지하고 있다는 것을 확인

03. 욕설탐지



결과



혐오와 중립이 많은 비율을 차지하고 있다는 것을 확인

03. 욕설탐지



결과 해석



분노/혐오/중립 감정에서 많은 욕설댓글 탐지

특히 정치/경제 카테고리의 혐오+분노의 비율이 60~80%

| | |
|---|----|
| 사람 새 귀 맞나? | 중립 |
| 여기 기사에 조선족중국인들 단체로 관광왔네. 헛소리말고 너네 연변으로 가라. 잊 | 중립 |
| 저놈도 취임식에 초대했었어? 거의 짐승급이던데...혹 이승만의 정치깡패 이정재 | 중립 |
| 강 사기꾼임. 원래 사기꾼이 사기쳐놓고 이말저말 하면서 시간끌거든. | 중립 |
| 석렬이 애비 일본장학생 1호는 어떻게 생각 하고?? | 중립 |
| 앞으로도 어쩔 수 없이 자의든 타의든 국민의 힘의 강세는 계속되겠구나. 극혐 정청 | 중립 |
| 바이든 한테 호구딜 당하는거 보더니 인니찌끄레기도 kf21호구딜 하러 오네 | 중립 |

분노/혐오 댓글 외 나머지 감정에서 욕설로 감지된 댓글 따로 확인!



카테고리별로

혐오(or 분노)의 비율 4~7%p 더욱 클 가능성 有
(특히 정치분야에서는 13% 추가탐지)

03. 욕설탐지



결과 해석



분노/혐오/중립 감정에서 많은 욕설댓글 탐지

특히 정치/경제 카테고리에는 혐오+분노의 비율이 60~80%

| | |
|---|----|
| 사람 새 귀 맞나? | 중립 |
| 여기 기사에 조선족중국인들 단체로 관광왔네. 헛소리말고 너네 연변으로 가라. 잊 | 중립 |
| 저놈도 취임식에 초대했었어? 거의 짐승급이던데...혹 이승만의 정치깡패 이정재 | 중립 |
| 강 사기꾼임. 원래 사기꾼이 사기쳐놓고 이말저말 하면서 시간끌거든. | 중립 |
| 석렬이 애비 일본장학생 1호는 어떻게 생각 하고?? | 중립 |
| 앞으로도 어쩔 수 없이 자의든 타의든 국민의 힘의 강세는 계속되겠구나. 극혐 정청 | 중립 |
| 바이든 한테 호구딜 당하는거 보더니 인니찌끄레기도 kf21호구딜 하러 오네 | 중립 |

분노/혐오 댓글 외 나머지 감정에서 욕설로 감지된 댓글 따로 확인!



카테고리별로

혐오(or 분노)의 비율 4~7%p 더욱 클 가능성 有
(특히 정치분야에서는 13% 추가탐지)

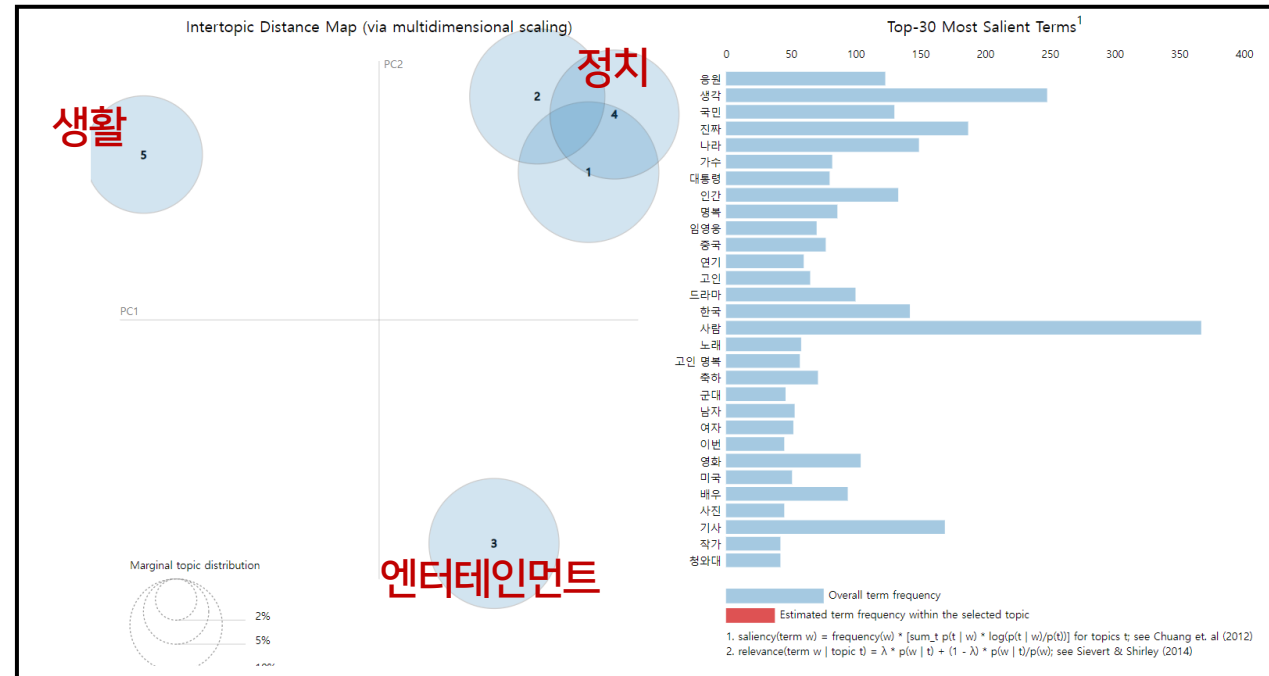


4. LDA



결과 해석

생활문화 전체



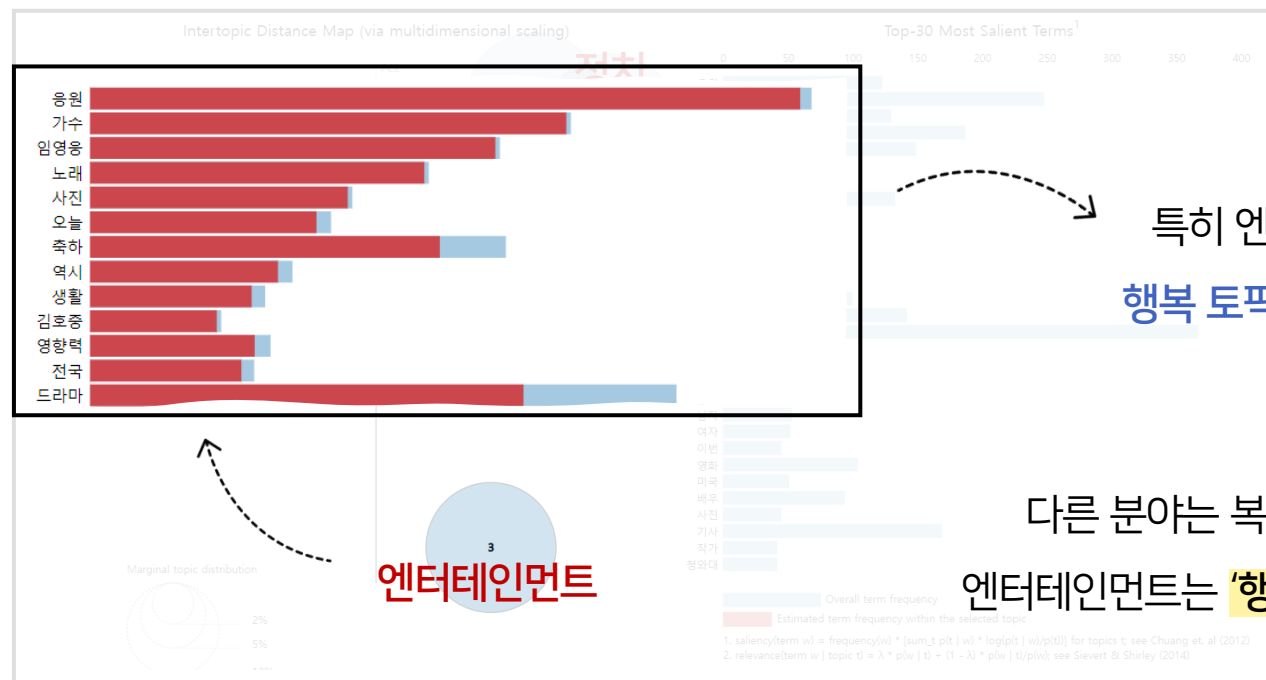
감정구분 없이 생활문화 전체적으로 토픽을 분석해봤을 때,
크게 세 분야(정치, 엔터테인먼트, 생활)로 구분

04. LDA



결과 해석

생활문화 전체



특히 엔터테인먼트 분야는
행복 토픽 키워드와 매우 유사

다른 분야는 복합적인 감정이 나타난 반면
엔터테인먼트는 '행복'이 우세한 것으로 해석 가능

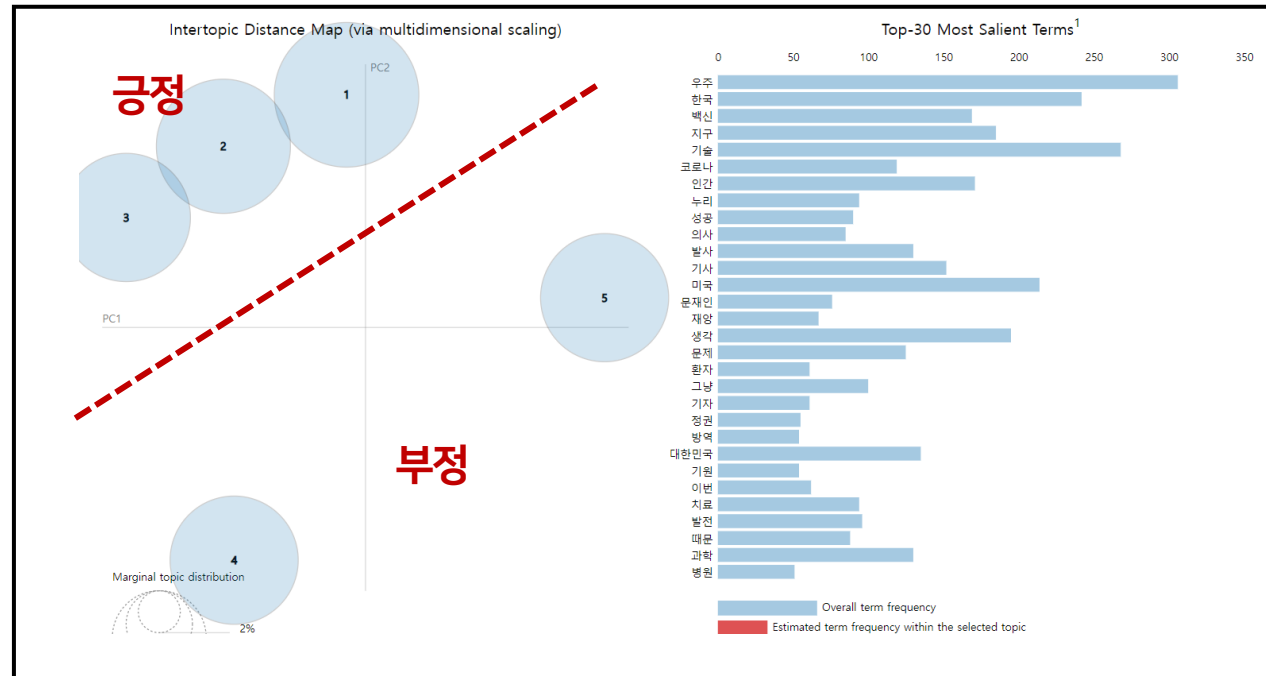
감정구분 없이 생활문화 전체적으로 토픽을 분석해봤을 때,
크게 세 분야(정치, 엔터테인먼트, 생활)로 구분

04. LDA



결과 해석

과학 전체



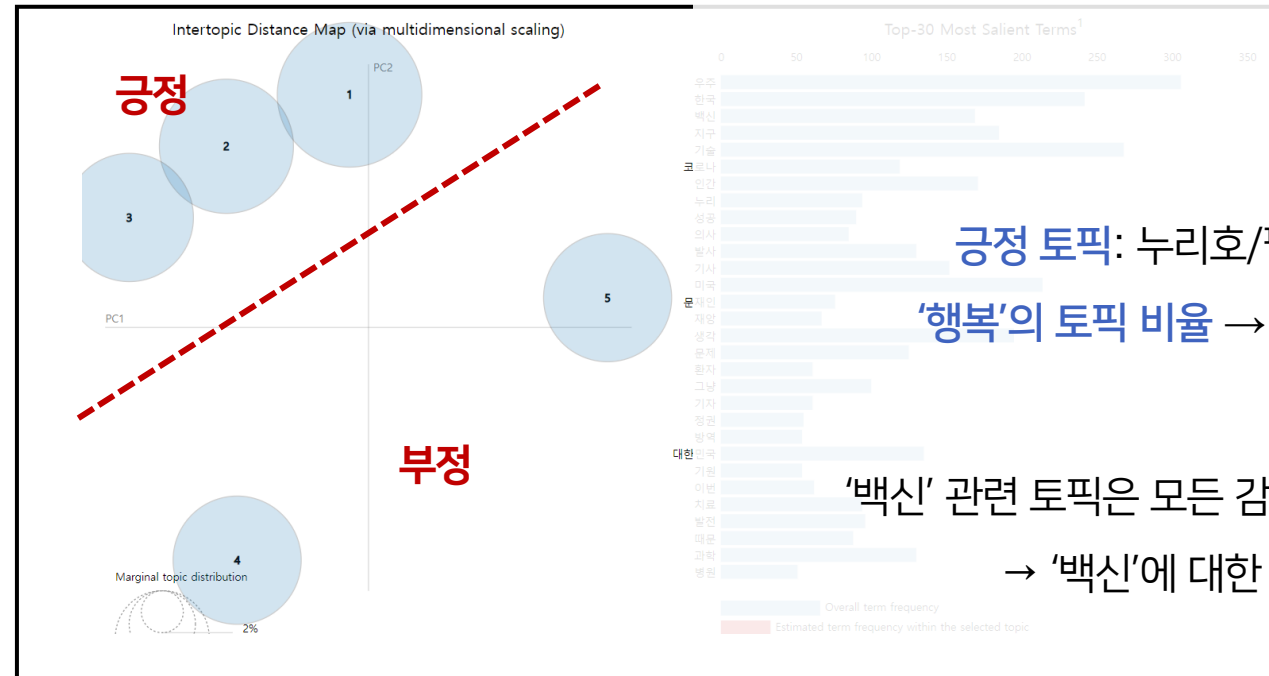
같은 방식으로 과학 분야 또한 분석
대각선 기준으로 크게 긍/부정으로 분류 가능

04. LDA



결과 해석

과학 전체



같은 방식으로 과학 분야 또한 분석
대각선 기준으로 크게 긍/부정으로 분류 가능



결과 분석



행복을 드러내는 양상

대부분 2개의 토픽을 가짐 → 축하, 응원, 감사 등 유사한 양상(키워드)을 보임



정치 관련 토픽

모든 카테고리에서 정치 관련 토픽이 존재 → 대부분 **부정적인(분노/혐오)** 반응.

욕설로 탐지되었던 단어들이 키워드로 선택되는 경우 多



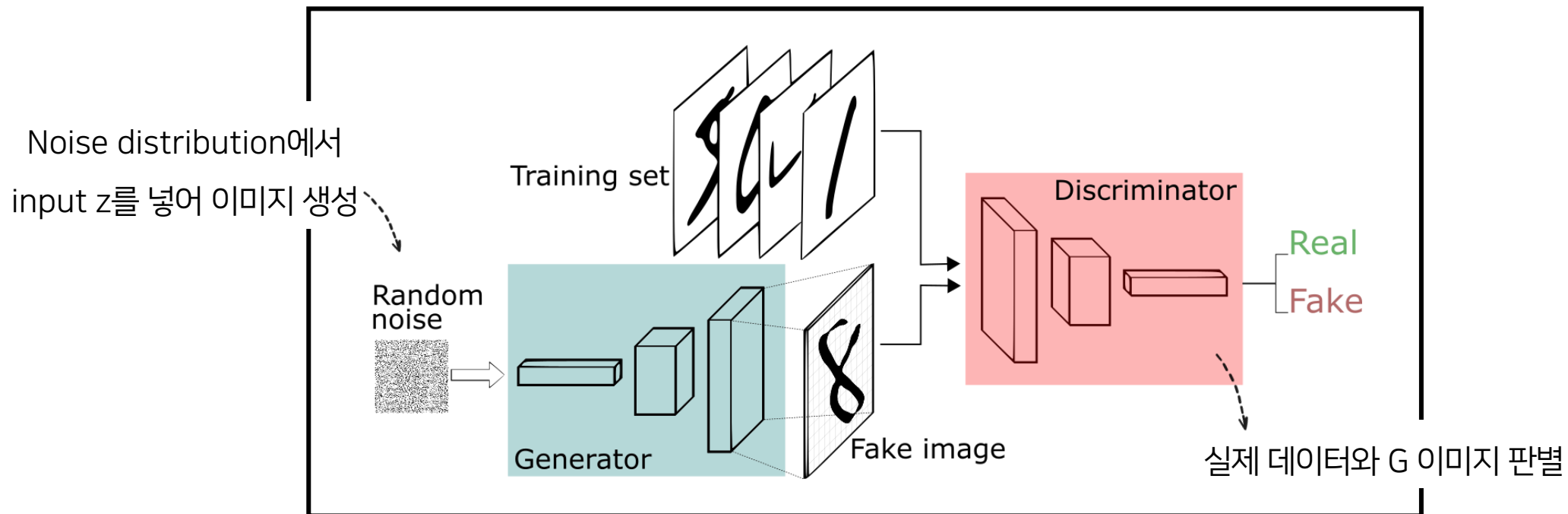
정치/경제/세계 카테고리: 정치 분야가 대부분의 토픽을 차지, 토픽의 구분이 명확 X

정치적 입장이 극명하게 다른 사용자들이 많음 → 같은 토픽임에도 단어 선택이 다름 → 구분이 모호



5. GAN

GAN



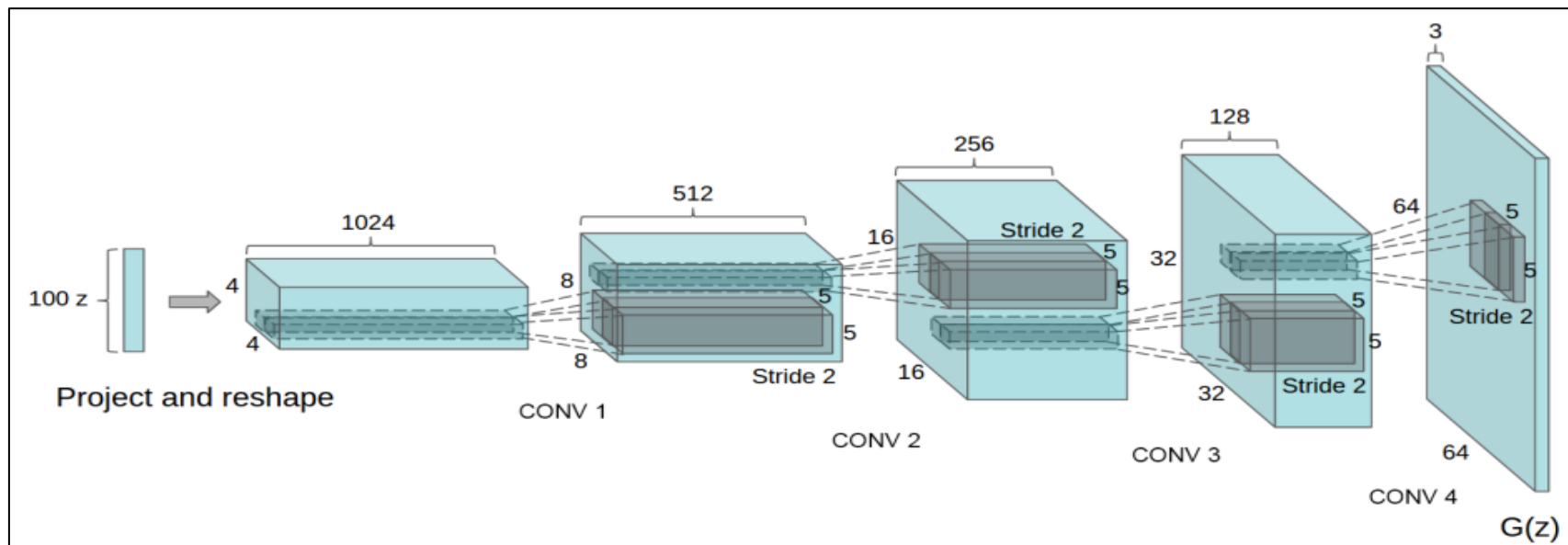
판별자 (Discriminator): 최대한 가짜/진짜를 잘 구별하기 위해 노력

생성자 (Generator): 최대한 진짜 같은 가짜 이미지를 만들어내려 노력

05. DCGAN



DCGAN



Deep **Convolutional** Generative Adversarial Network

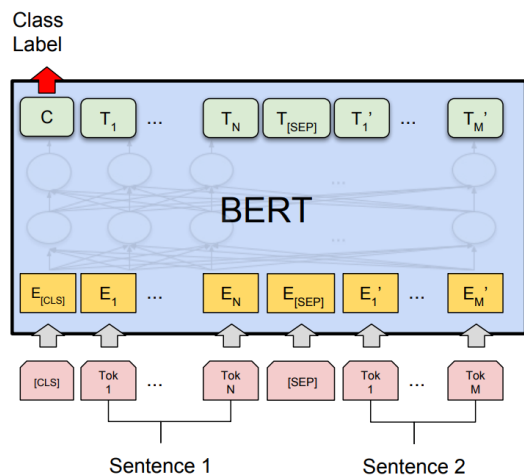
↓
생성자에 Convolutional 구조를 녹여냄

05. DCGAN 구현

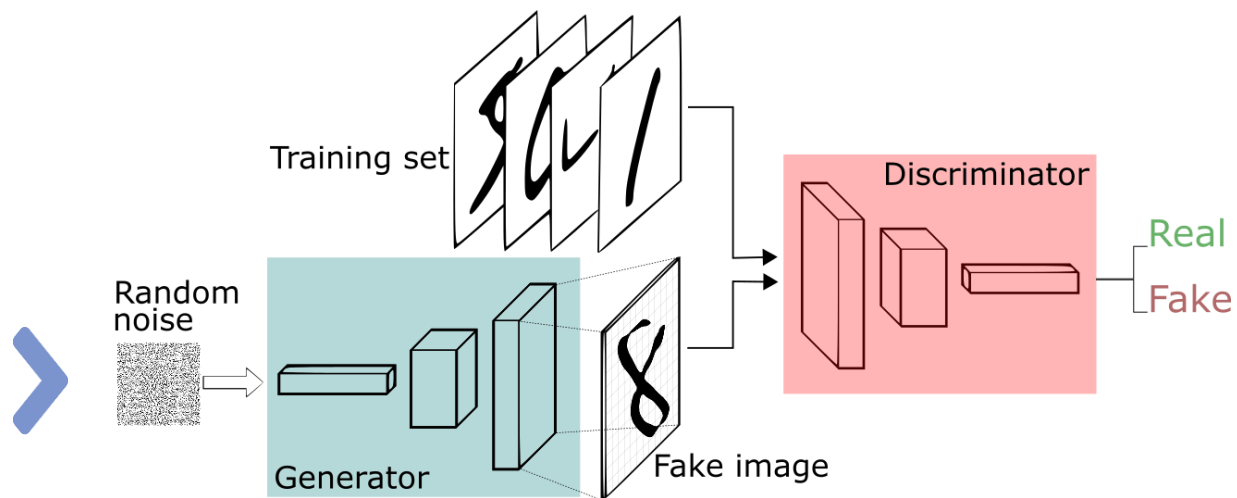


진행 과정

그렇다면 우리는...?



코버트를 이용해 댓글의 **임베딩 벡터**를 구하고
그 벡터값을 노이즈 대신 넣음

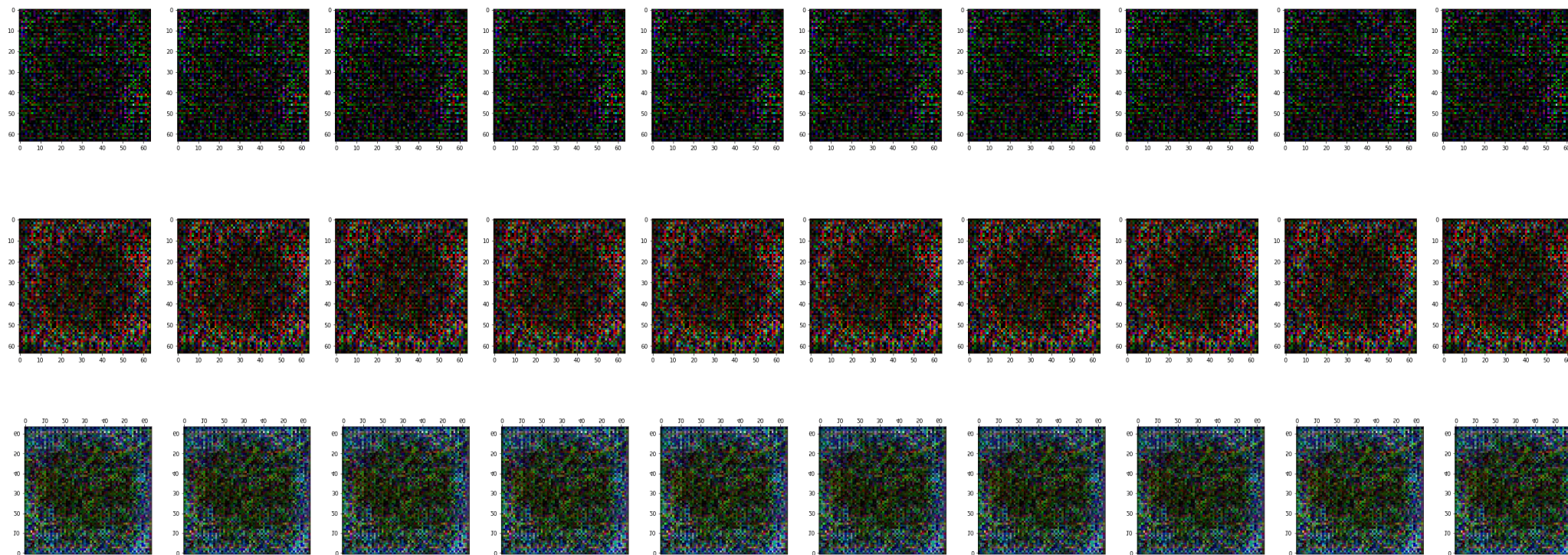


감정으로 라벨링한 댓글의 **임베딩 벡터**로 학습시켜
생성자가 '**감정**'이라는 **잠재요소**를 파악할 것으로 기대!

05. Loss를 보완한 GAN



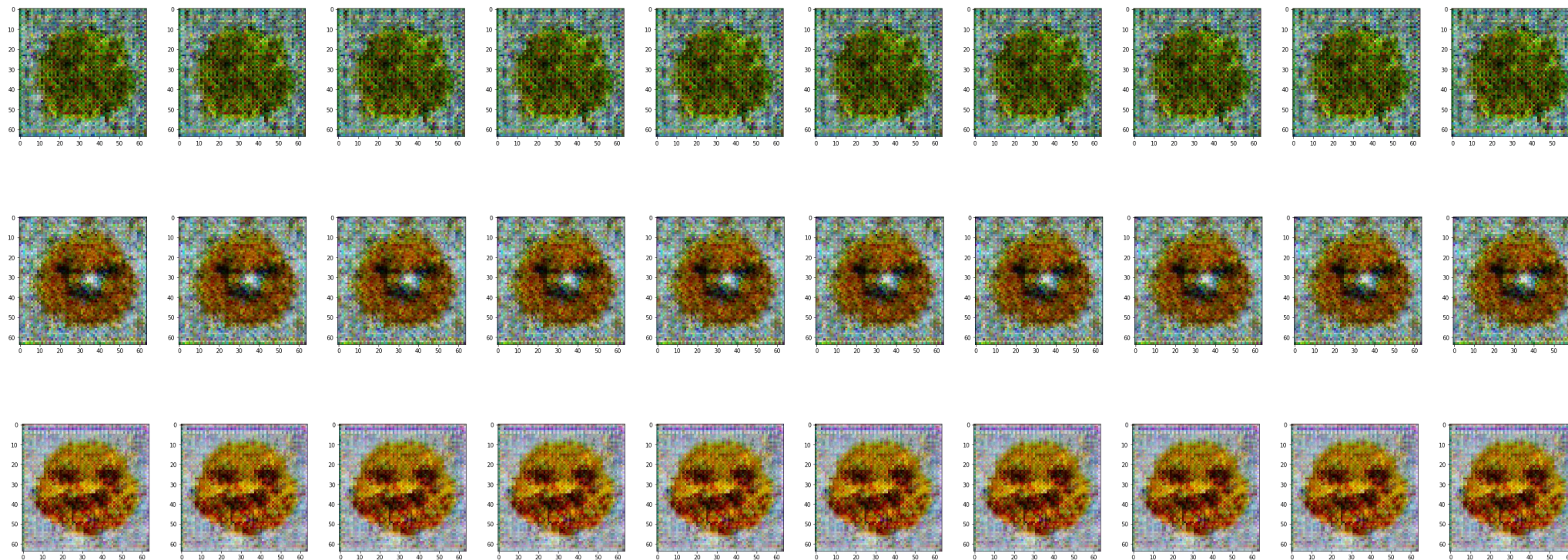
생성과정



05. Loss를 보완한 GAN



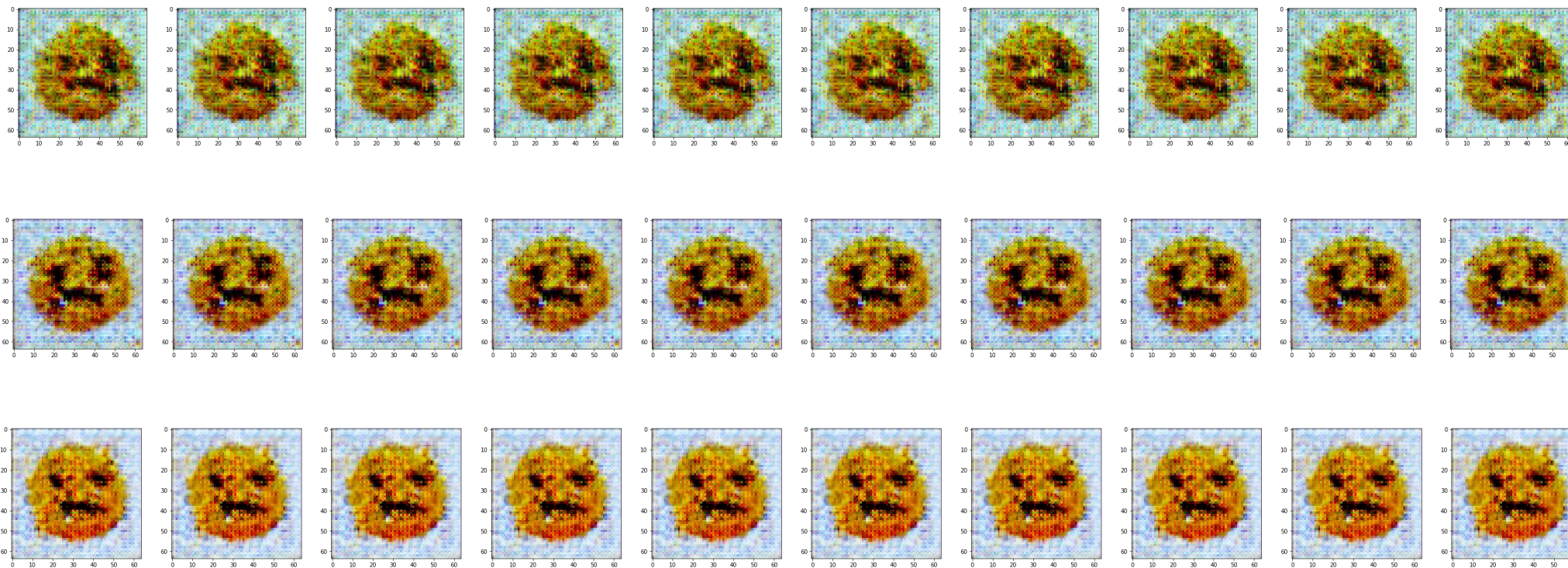
생성과정



05. Loss를 보완한 GAN



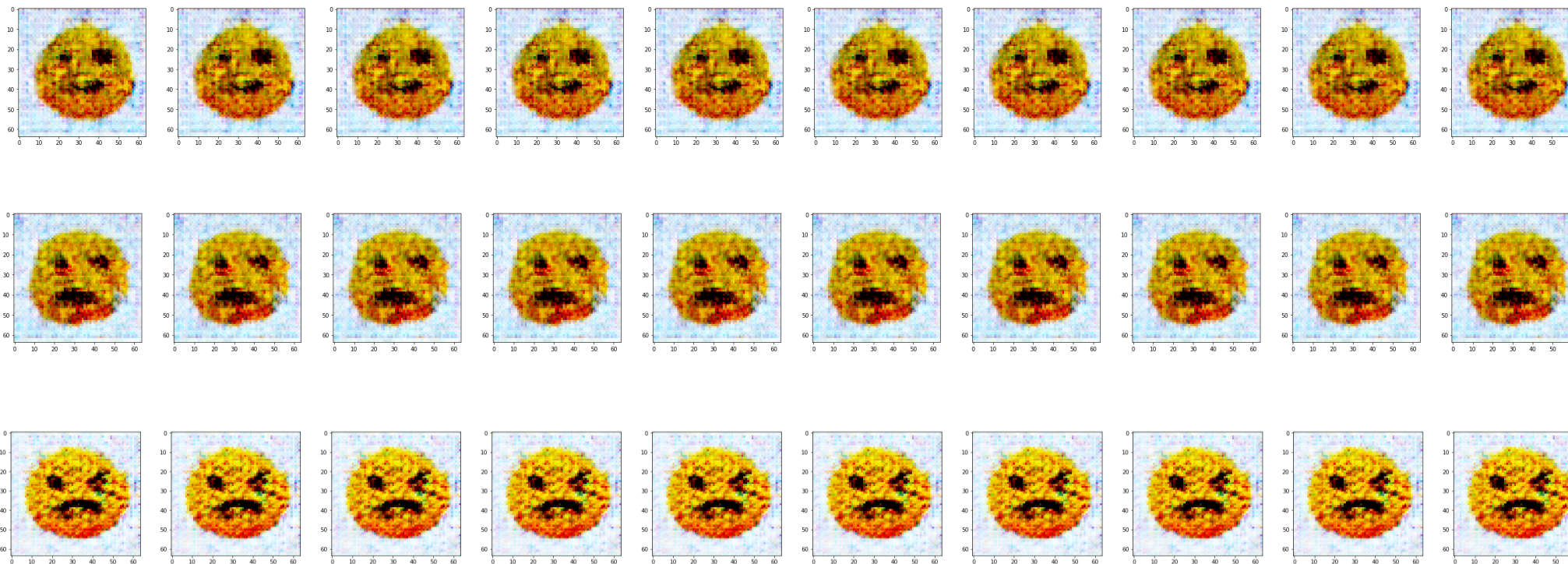
생성과정



05. Loss를 보완한 GAN



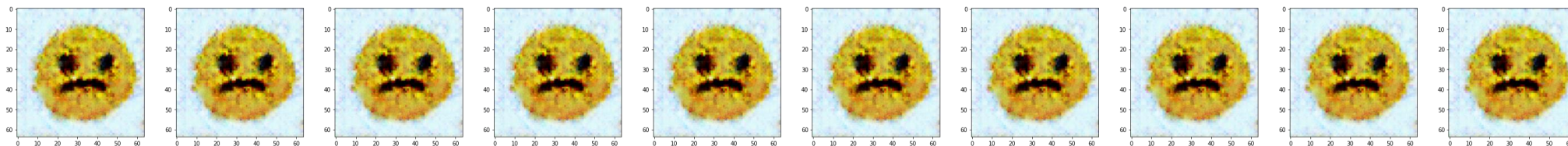
생성과정



05. Loss를 보완한 GAN



생성과정



Gif로 보고 싶으신 분은 따로 연락주세요...

05. Loss를 보완한 GAN



적용



코플플 쓰면서 컴퓨팅 단위 깎일 때마다 제 수명도 같이 깎이는 기분이었습니다.
GPU 할당 서비스 추가도 안 해주는 정 없는 *글 당분간 불매합니다;;



그누무 풀뿌리 뿌리채 뽑고 싶네.
세금먹는 블랙홀



#번 찍은 꿀통들 때문에 별 그지같은게 호와호식하며
땡땡거리고 산다. 이나라 아직 멀었다.



이거 화나서 콧김 뿜는 것으로 예상..

개판이지. Y이는 또 모르는일이겠지.



어디쓰레기미국태풍.이나



공소시효 없애라 죄지은 놈은 죽기 전까지 찾아서 처벌해야 된다

라벨링 비율이 높은 분노/혐오의 감정을 담은 댓글은 감정인식 & 이미지 생성 모두 상대적으로 안정적

05. Loss를 보완한 GAN



적용

: 행복



너무 좋아 진짜 최고야



멋지십니다 굳건하세요 응원합니다



정동원 봉화 은어축제 축하합니다



최고리그 득점왕!!!!

행복 또한 감정인식 & 이미지 생성 모두 **상대적으로 안정적**

05. Loss를 보완한 GAN



적용

: 슬픔



저를 매우 치세요 늦을 것 같습니다.. 달려가겠습니다.. 죄송합니다 ㅜㅜ



삼가 고인의 명복을 빕니다..



지난 5년간 무슨일이 있었기에 나라가 이모양이 된거냐 ㅜㅜ
온갖 정부 부서가 다 먹통이네...



다행입니다..... ㅜㅜㅜ 다행이에요

슬픔의 경우 감정은 잘 인식하지만
이미지 생성에서 불안정한 모습을 보임

05. Loss를 보완한 GAN



적용

: 그 외 감정들



집에 가고 싶다...



저는 소식좌가 아닙니다... 조금씩 자주 먹는 겁니다...



더운데 감사합니다~



컴퓨터 로그기록 뒤지면 되지
밤시간인데



물만 주면 될. 일을..



과학 방역? 각자도생



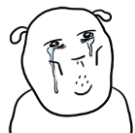
삼가 세분의 명복을 빕니다
평안히 영면 하소서

중립 / 중의적 문장 / 해석하기 어려운 문장 / 오타가 많은 문장 등
감정을 잘못 나타내거나 알 수 없는 표정을 나타내기도 함

06. 한계 및 의의



한계와 의의



한계

원래 목표는 벡터 연산을 통해 감정과 댓글 내용을
같이 나타내는 이모티콘을 만드는 것이었지만
시간이 부족했습니다...

식별 가능한 이모티콘이 생성되었다



의의

GAN에 대해 더 자세히 알 수 있어서 좋았다

판별자를 조정해가며 개념적으로만 알고 있던 적대적
학습의 알고리즘의 진행과정을 눈으로 확인할 수 있었다

식별 가능한 이모티콘이 생성되었다