

## 텍스트\_추가 전처리.txt

✓ 데이터 양이 많지 않아 전체 데이터 사용

전체 단어 집합 크기 = 27480

```
0          저녁
1      콩국수 드실 분
2  현재 인원 강지원 한상현
3          오우
4      리액션 봇
...
53196      이모티콘
53197      고대 에타 에 올리자
53198      나쁜 서연
53199      과제 얼른 끝내구 황금 연휴 누려
53200      누나 도 연휴 잘 보내 새용
```



단어가 비효율적으로 너무 많아서  
전체 단어를 사전으로 이용 불가능!!



사용 빈도가 높은 단어를 쓰자!

## 텍스트\_추가 전처리.txt

### 1. Float Encoding

각 데이터를 정수 인코딩 후 **하나의 문장을 여러 문장으로 분해**하여 train data 구성  
=> 다음에 나올 단어를 예측하기 위해 이전에 나온 단어들을 참고해야 하므로

example

X	y
콩국수	드실
콩국수 드실	분
현재	인원
현재 인원	강지원
고대	에타
고대 에타	에
고대 에타 에	올리자

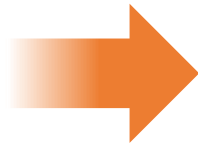
Float encoded Dataset

```
[[1112, 1139],  
 [1112, 1139, 1432],  
 [166, 175],  
 [108, 58],  
 [108, 58, 626],  
 [2, 154],  
 [2, 154, 305],  
 [2, 154, 305, 154],  
 [2, 154, 305, 154, 961]]
```

텍스트\_추가 전처리.txt

## 2. Padding

```
[[1112, 1139],  
 [1112, 1139, 1432],  
 [166, 175],  
 [108, 58],  
 [108, 58, 626],  
 [2, 154],  
 [2, 154, 305],  
 [2, 154, 305, 154],  
 [2, 154, 305, 154, 961]]
```



X 변수							Y 변수
[	0,	0,	0,	...,	0,	1112,	1139]
[	0,	0,	0,	...,	1112,	1139,	1432]
[	0,	0,	0,	...,	0,	166,	175]
...,							
[	0,	0,	0,	...,	584,	9,	1882]
[	0,	0,	0,	...,	9,	1882,	33]
[	0,	0,	0,	...,	1882,	33,	682]

## 이미지\_추가 전처리\_Labelling

	Bangs	Black_Hair	Brown_Hair	Chubby	Eyeglasses	Heavy_Makeup	Male
1.jpg	1	1	-1	1	-1	-1	1
2.jpg	-1	1	-1	1	1	-1	1
3.jpg	-1	1	-1	-1	-1	-1	1
4.jpg	-1	1	-1	-1	-1	-1	1
5.jpg	-1	1	-1	-1	-1	-1	1
6.jpg	-1	1	-1	-1	-1	-1	1
7.jpg	-1	1	-1	-1	-1	-1	1
8.jpg	-1	1	-1	-1	-1	-1	1
9.jpg	-1	1	-1	-1	-1	-1	1
10.jpg	-1	1	-1	-1	-1	-1	1
11.jpg	-1	1	-1	-1	-1	-1	1
12.jpg	-1	1	-1	-1	-1	-1	1
13.jpg	-1	1	-1	-1	-1	-1	1
14.jpg	-1	1	-1	-1	-1	-1	1
15.jpg	-1	1	-1	-1	-1	-1	1
16.jpg	-1	1	-1	-1	1	-1	1
17.jpg	-1	1	-1	-1	1	-1	1
18.jpg	-1	1	-1	-1	1	-1	1
19.jpg	-1	1	-1	-1	-1	-1	1

라벨링 완성 (최종 변수 15개)

(Bangs, Black\_hair, Brown\_hair, Chubby, Eyeglasses, Heavy\_makeup, Male, Mouth\_Slightly\_Open, Narrow\_Eyes, Pale\_Skin, Rosy\_Cheeks, Smiling, Wearing\_Earrings, Wearing\_Hat, Wearing\_Lipstick)

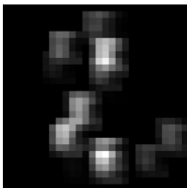
# 생성 모델\_VAE



인코더

[ 1.5, -0.5 ] 잠재 벡터

디코더



## Auto encoder

입력 데이터와 재구성 데이터 사이의 손실을 최소화하는  
인코더와 디코더의 가중치를 찾으며 훈련!

Encoder

고차원 입력 데이터를  
저차원의 잠재 벡터로  
압축해서 구성

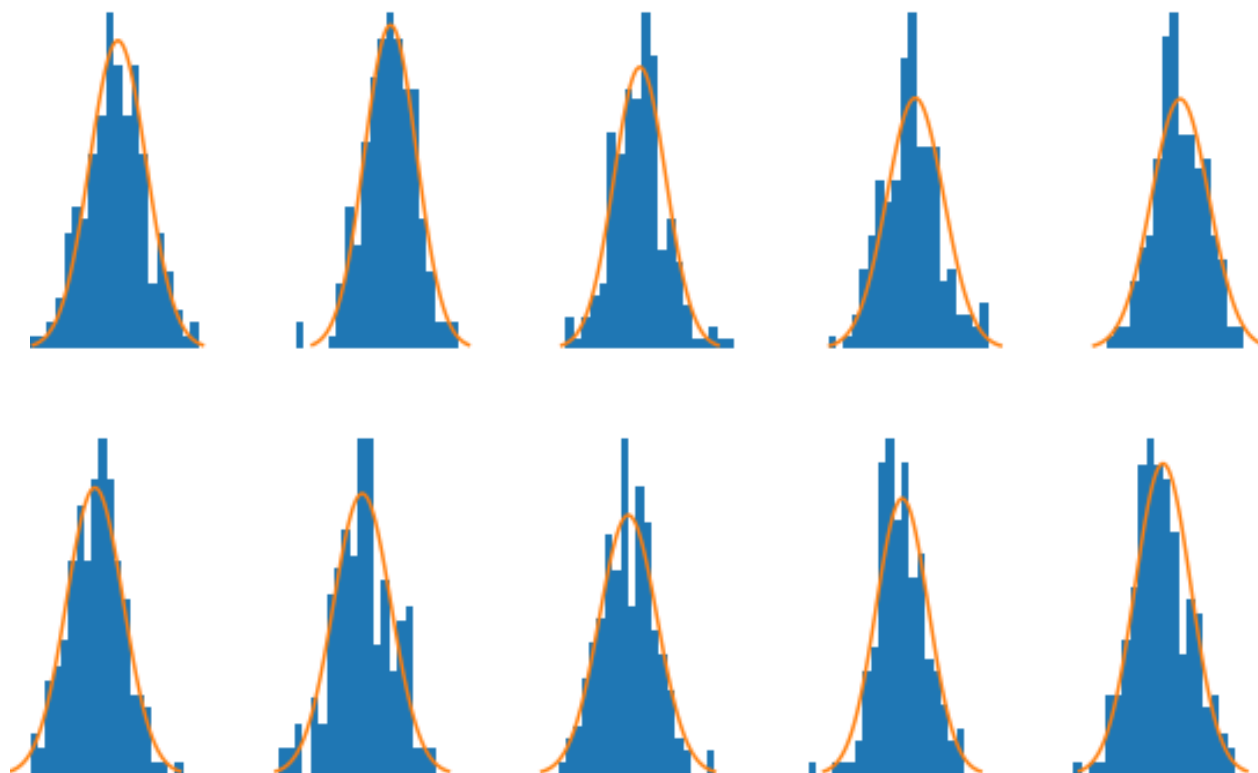
Decoder

주어진 잠재 벡터를  
원본 차원으로 재구성

## 생성 모델\_VAE



모델을 완성했으니 다시 표준정규분포를 잘 따르는 지 확인해보자



## VAE 생성모델



확실히 선명하고 진짜 사람 같은 이미지가 많이 생성되는구나!

Epoch  
400~500



Epoch  
500~700



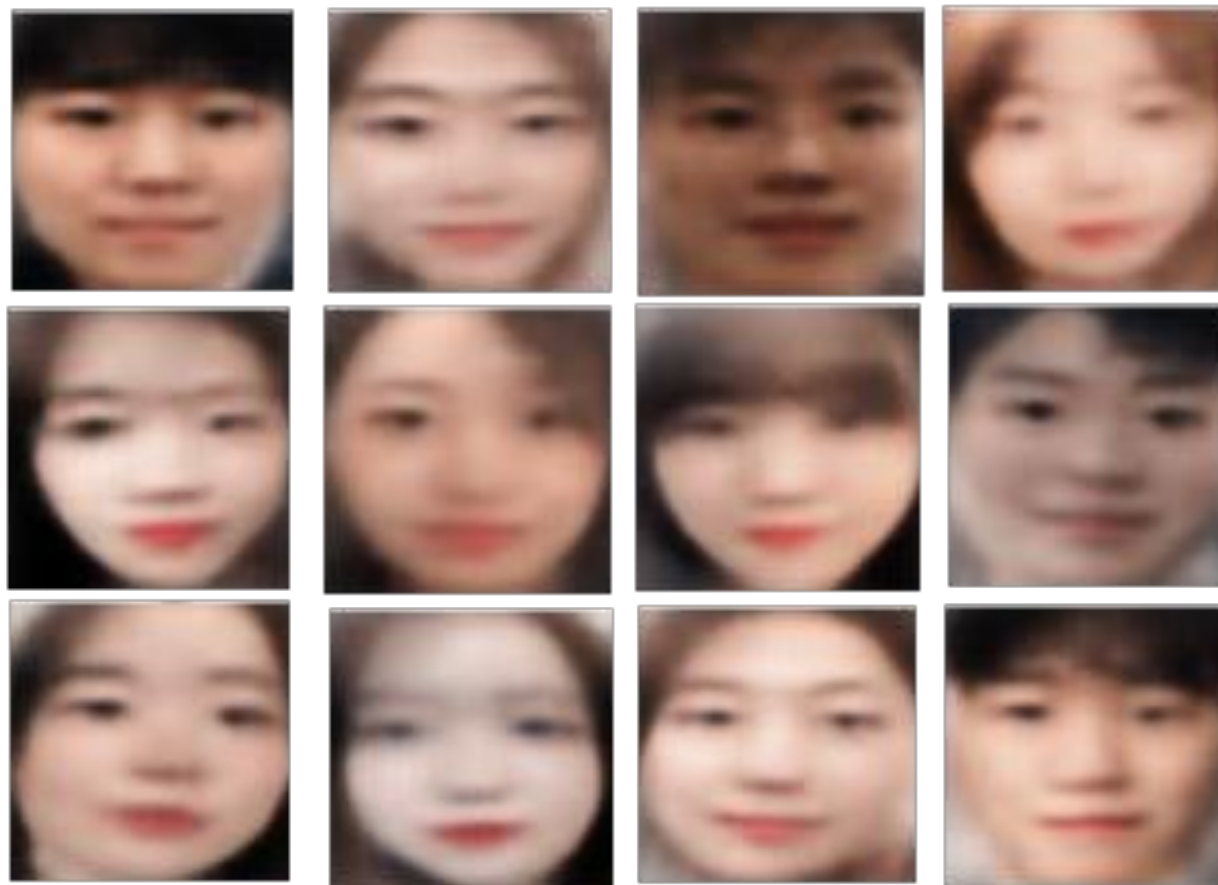
Epoch  
700~800



## 생성 모델\_VAE

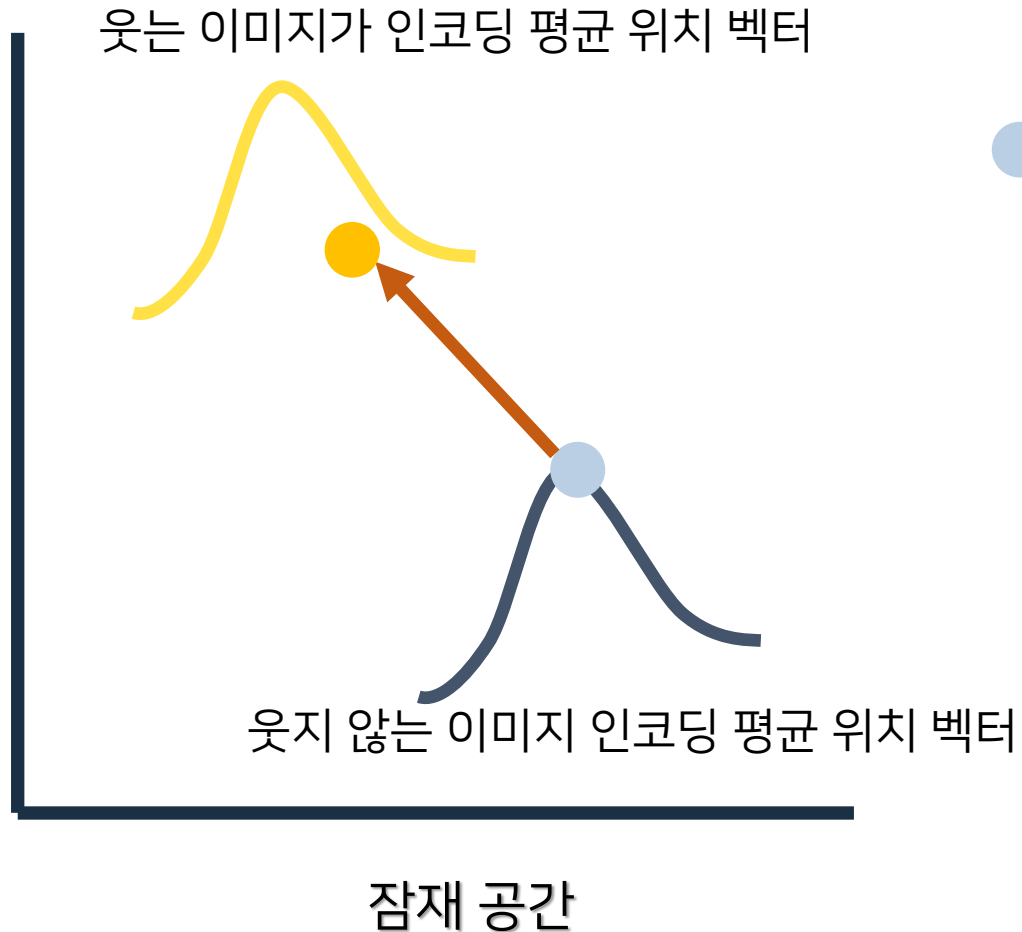


최종 모델로 이미지를 생성해보자





원본 이미지의 특징을 바꿔보자!!



● [ 원본 이미지의 인코딩 벡터 ]

+  [ 웃음으로 향하는 벡터 ]

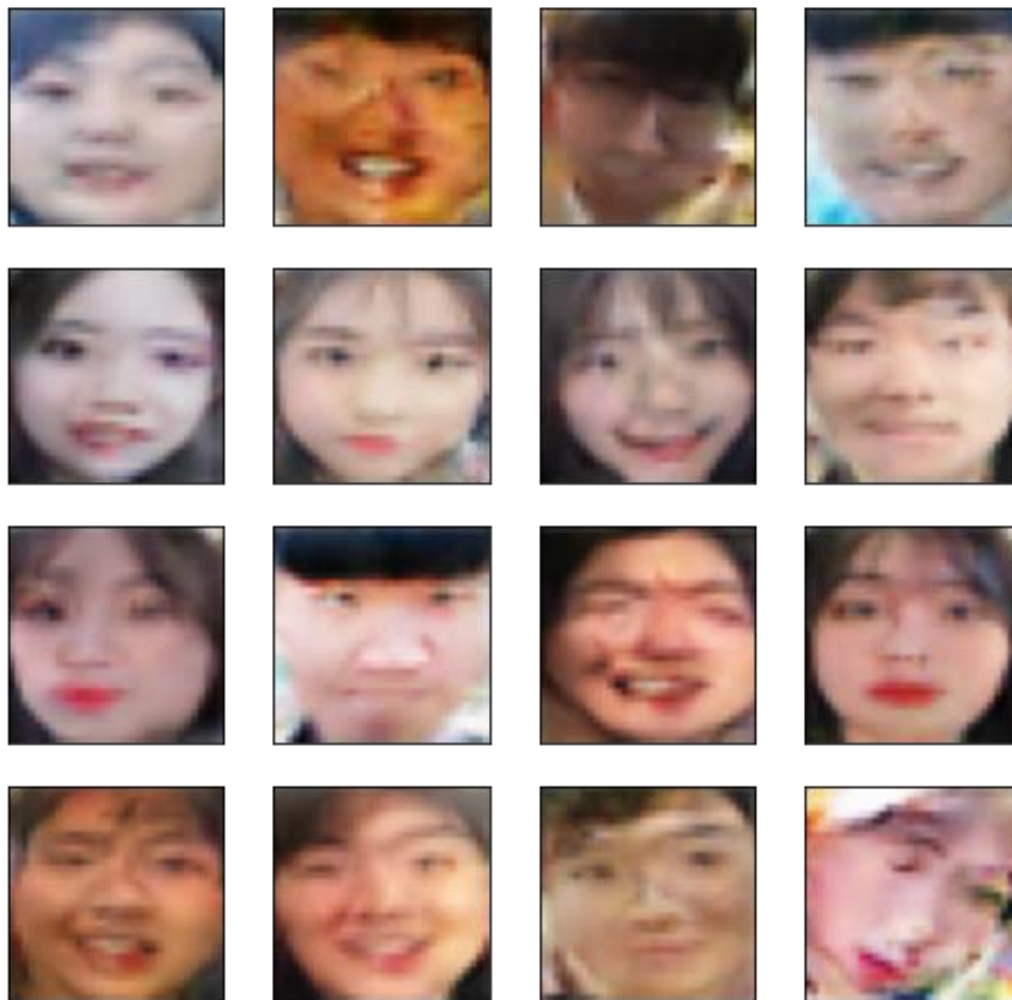
= ● [ 웃음 벡터가 더해진 새로운 포인트 ]

새로운 포인트를 디코딩하여 웃는 표정의 이미지 생성

## 모델링\_DCGAN\_#3



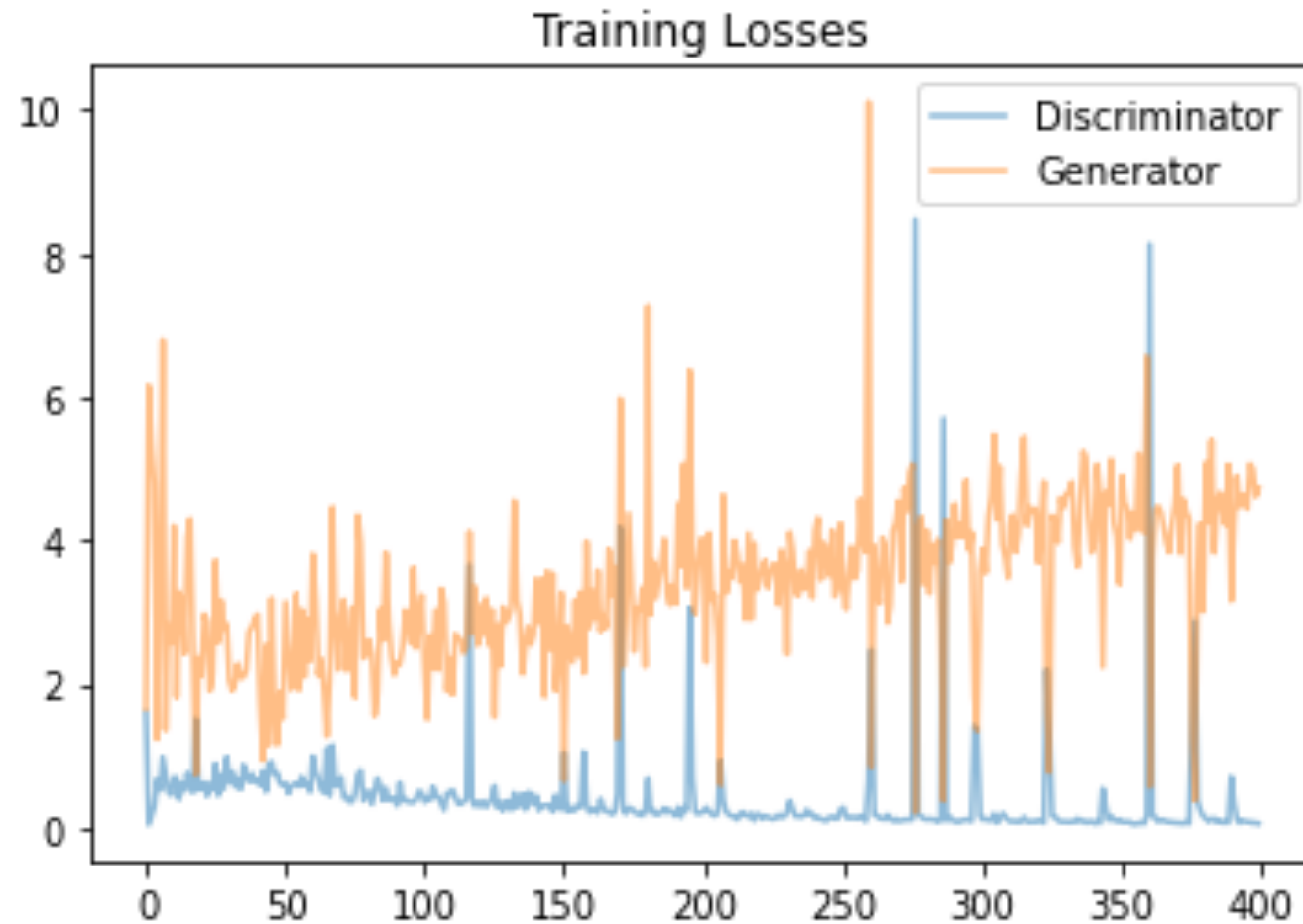
Epoch 400번 으로 늘렸다! (8시간이나 걸리다니ㅠ)



## 모델링\_DCGAN\_#3



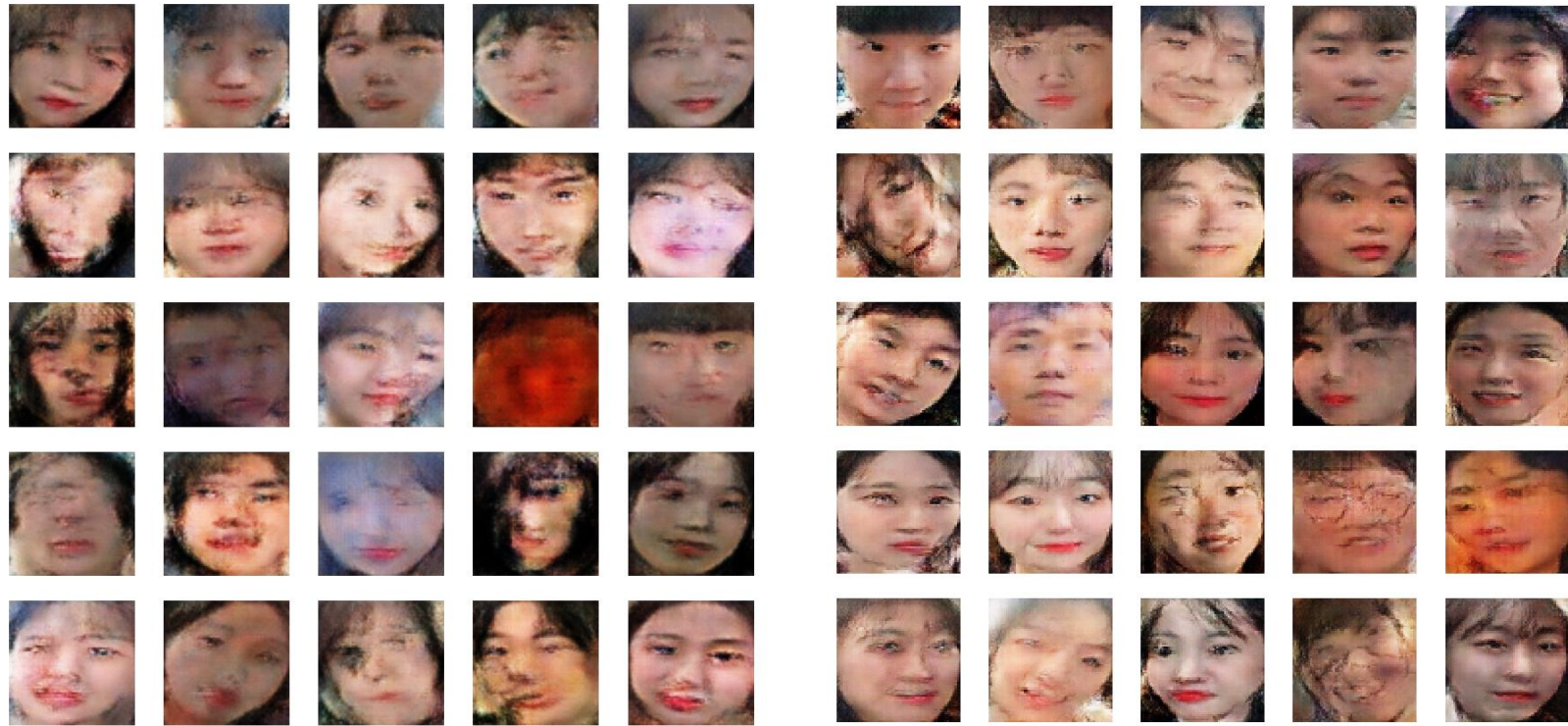
Discriminator loss는 안정적이고, Generator loss는 더 높아졌넹 이전보다 더 좋은 결과구나!! 신낭당!!!!



# 모델링\_WGAN-GP\_#3



장장 하루 반에 걸친 Epoch 10000...!  
이 정도가 데이터의 한계인 듯...!



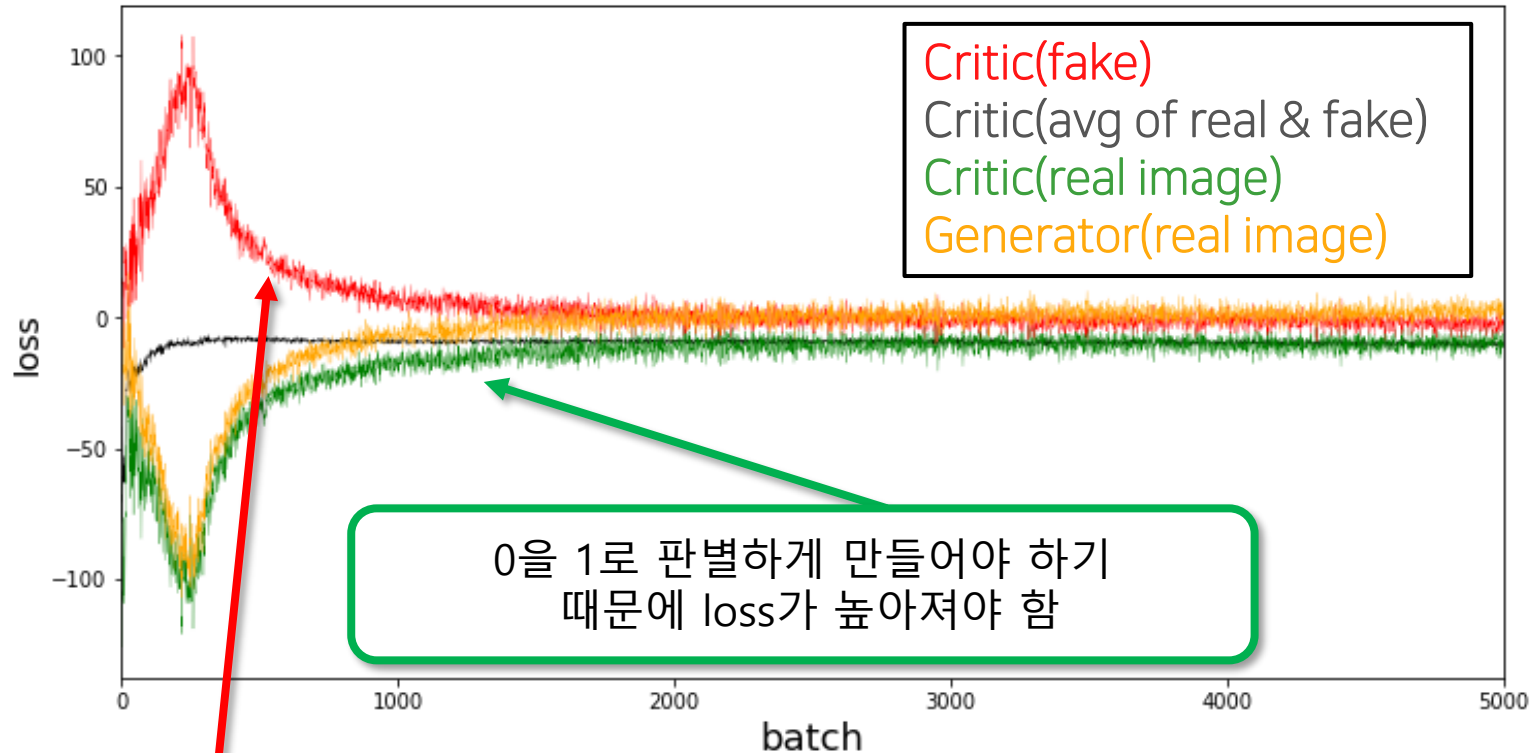
Epoch=300

Epoch=10000

## 모델링\_WGAN-GP\_#3



매우 이상적인 학습 곡선!



0을 1로 판별하게 만들어야 하기  
때문에 loss가 높아져야 함

진짜/가짜로 판별할 확률이 0.5(강 막 찍는 수준;)가  
되어야 하기 때문에 학습이 길어질수록 평행/수렴함



# LSTM vs GRU

Bi-LSTM & Bi-GRU

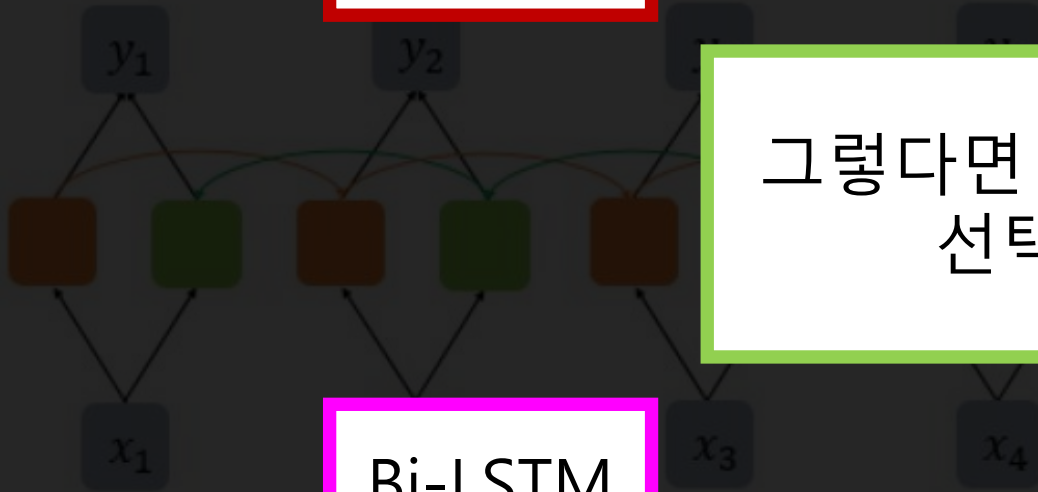
LSTM

GRU

그렇다면 어떤 모델을  
선택할까?

Bi-LSTM

Bi-GRU



특정 시점의 값을 예측하기 위해  
시점 데이터 + 이후 시점 데이터  
메모리 셀 사용  
메모리 셀 : 앞 시점의 은닉상태 전달  
연두색 메모리 셀 : 뒤 시점의 은닉상태 전달  
=> 두 개의 값을 예측하는데 사용

# Bi-GRU

Data	단어 집합 개 수	Epoch	Batch_size
Okt	2000	400	128

## Bi-GRU 결과

```
print(sentence_generation(model_gru, t, '다빈', 12))
```

다빈 오빠 도 갸륵 으로 조만간 예정 이니 걱정 하지 말라구 순 일

```
print(sentence_generation(model_gru, t, '윤혜', 5))
```

윤혜 야 생일 축하 해 🎉

```
print(sentence_generation(model_gru, t, '정훈', 9))
```

정훈 오빠 가 한 거 아니야 저 거 동영 이

```
print(sentence_generation(model_gru, t, '아니', 2))
```

아니 근데 도대체



세계관 최강자  
Bi-GRU 수준 실화냐...

```
print(sentence_generation(model_gru, t, '피셋', 7))
```

피셋 의 희망 것 이 없는 것 같은데

```
print(sentence_generation(model_gru, t, '밥', 4))
```

밥 먹으러 어디 쪽 으로

```
print(sentence_generation(model_gru, t, '술', 6))
```

술 먹는 거 저 거 뭐 여

```
print(sentence_generation(model_gru, t, '회식', 4))
```

회식 함 조지 자 ❤️

```
print(sentence_generation(model_gru, t, '현지', 8))
```

현지 생일 축하 해 현지 야 진짜 축하 해

# 26기 피셋이들 ><



다빈 오빠 도 갠톡 으로 조만간 예정 이니 걱정 하지 말라구 순 일

흐뭇하구만,,,,,



윤혜 야 생일 축하 해 🎂



정훈 오빠 가 한 거 아니야 저 거 동영 이



회식 함 조치 자 ❤️



술 먹는 거



현지 생일 축하 해 현지 야 진짜 축하 해



밥 먹으러 어디 쪽 으로