

전 은 하

PORTFOLIO

01

자기소개

1

02

마케팅기획

2

- 2.1 핑크플레이콘서트
 토넷비누 키트 제작

03

세일즈포스(Saas)

3

- 3.1 세일즈 포스 운영 관리
 및 영상제작

04

AI빅데이터MBA 프로젝트

- 4.1 R활용한 다중회귀분석
4.2 캐글 -머신러닝 예측
4.3 캐글 - 딥러닝 감성분석



업무 전문성을 갖고 꾸준히 최선을 다하는 인재

Rather be death than cool

이름 : 전은하

1생년월일 : 1989.02.08(만32세)

거주지 : 안산

경력 연차 : 7년 11개월

특기 : 사진, 독서

취미 : 영상, 운동

입사가능시기 : 8월

010 2561 3708

eunhaj02@naver.com

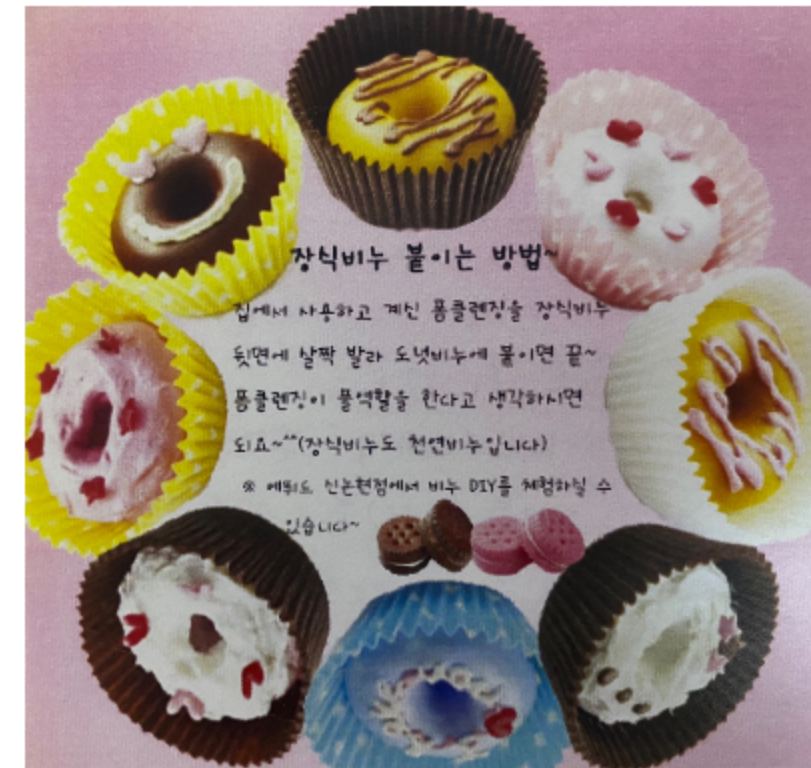


핑크 플레이 콘서트 도넛키트 제작

2012.04 ~ 2012.08

(주)에뛰드 Marketing Innovation 인턴
장애우 친구들이 만든 비누가 새로운 제품
으로 창출되는 과정을 체험

- 에뛰드에서 '비누'라는 개념이 하나의 '스스로 만들어나가는 피부', 셀프 화장의 이미지로 바뀌는 과정을 배움
- 논현점 인테리어 설치에 도움
- 핑크플레이콘서트 부스 설치 및 판매



리플렛 제작

ETUDE HOUSE

© 2018 ETUDE HOUSE. All rights reserved.



세일즈포스(Saas) 활성화 방안 기획

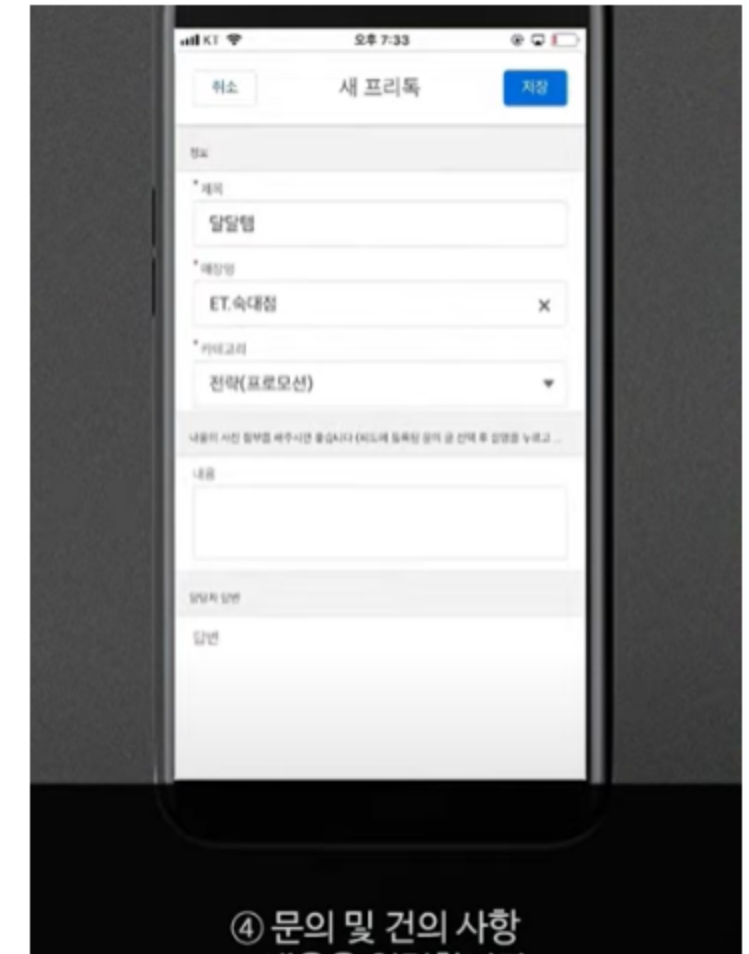
2017-08-01 ~ 2018-12-23

(주)에뛰드 리테일운영팀 사내 프로젝트(CRM)

Manual Video



카테고리 메뉴 소개



게시글 올리는 방법 소개

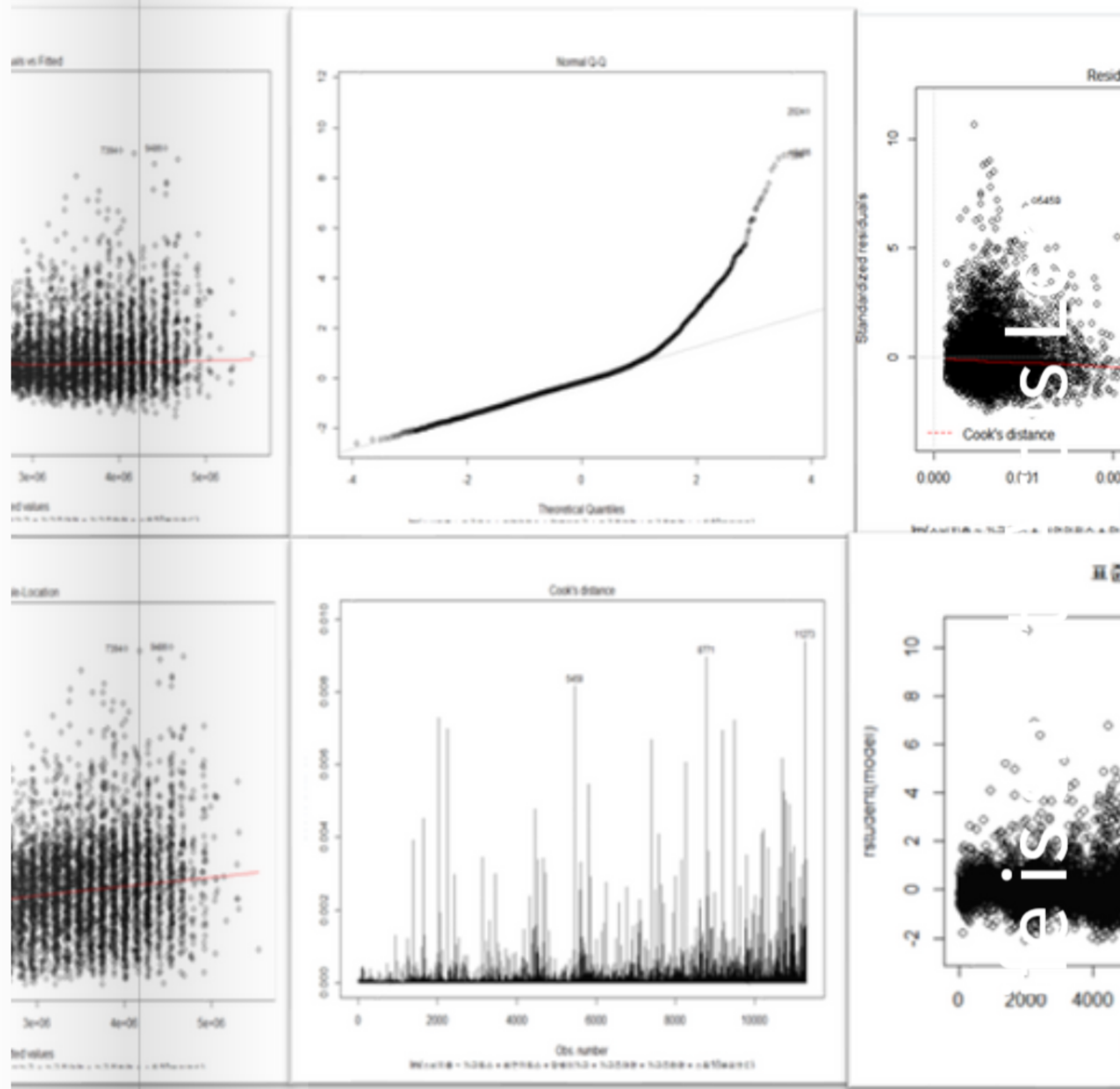
- 가맹점 계정 관리 및 업무효율화를 위한 UI화면 개선
- AS신청율 개선
- 가맹점주 및 협력사 대상 세일즈포스 교육 및 영상 제작
- 영업담당자 및 파트너사 직원의 현장방문율(매장/시장조사 월 1회 방문 지표) 개선
- 체류시간, 방문율, 현장방문지표 대시보드, GPS 오류 개선 등을 활용하여 영업담당자의 현장방문율 10% 제고
- 표준단가표, 알림 서비스 제공, 결재단계 최소화를 통한 업무 효율성 증대

R을 활용한 다중회귀분석

2019-12-04 ~ 2019-12-21

국민대학교 경영대학원 AI빅데이터MBA

- R 및 엑셀을 통한 자료 준비 및 팀 발표 수행
- 통계청 데이터를 가지고 R을 활용하여 회귀분석
- 데이터 시각화를 통한 발표자료 준비 및 발표
- 가구의 특성이 가계의 소비지출에 미치는 영향 관한 분석
- 가구 특성 중 6가지 연속형 변수들이 소비지출과 높은 상관관계를 가지고 있다는 가설 검증



Public Leaderboard

Private Leaderboard

The private leaderboard is calculated with approximately 60% of the test data.

This competition has completed. This leaderboard reflects the final standings.

Refresh

#	Δpub	Team Name	Notebook	Team Members	Score	Entries	Like
1	—	U2019548			0.94772	14	8r
2	—	HeechulYang			0.94597	25	8r
3	—	U2019515			0.94268	4	8r
4	—	bae jihoon			0.93465	15	8r
5	—	U2019518			0.92007	15	8r
6	—	29me			0.91471	6	8r
7	▲ 1	eunha jeon			0.91354	10	8r
8	▼ 1	jeonghwaaa			0.91309	22	8r
9	—	sm0507.Lee			0.91261	10	8r
10	—	Eunsil Lee			0.89309	63	8r

캐글(KAGGLE) Competition

2020-01 ~ 2020-02

국민대학교 경영대학원 AI빅데이터MBA

* 머신러닝(계절)기말 Competition

* 호텔 수요 예측을 위한 취소 거래 분류 모형 구축

- 전처리, 분류, 앙상블
- 파이프라인 구축

```
type = (book.hotel_type == 'R').astype('int')
is_refund = (book.deposit == 'Non Refund').astype('int')
customer_type = book.customer_type.isin(['Contract', 'Group']).astype('int')
book_dummies = pd.get_dummies(data=book, columns=['year', 'month', 'meal', 'book_method'])
```

```
)

('book_train_new.csv')
loaded
['is_canceled'], axis='columns')
y_train, y_test = train_test_split(x, y, stratify=y, test_size=0.3)
```

```
rsClassifier()
Regression(solver='sag')
restClassifier()
```

```
x, y_train)
x, y_train)
x, y_train)
```

```
#python37\lib\site-packages\sklearn\linear_model\sag.py:330: ConvergenceWarning: The max_iter was reached
without converge
not converge", ConvergenceWarning)
```

```
sifier(bootstrap=True, ccp_alpha=0.0, class_weight=None,
       criterion='gini', max_depth=None, max_features='auto',
       max_leaf_nodes=None, max_samples=None,
       min_impurity_decrease=0.0, min_impurity_split=None,
       min_samples_leaf=1, min_samples_split=2,
       min_weight_fraction_leaf=0.0, n_estimators=100,
       n_jobs=None, oob_score=False, random_state=None,
       verbose=0, warm_start=False)
```

- 각종 keras 창시자로 부터 배우는 Tensorflow 2.0 + keras특강, SK AC ADEMY 강의 참고하여 공부
- Stratified k-fold sample submission 사용하려다가 성능이 나아지지 않아 원본 데이터 제출
- 원핫인코딩, min-max scaler, drop-out 사용하여 성능 높이기 시도

Kaggle Prediction Competition							
2020-1 DeepLearning Final							
감성분석							
40 teams · 4 months ago							
Overview Data Notebooks Discussion Leaderboard Rules Team My Submissions Late Submission							
Public Leaderboard Private Leaderboard							
The private leaderboard is calculated with approximately 50% of the test data. This competition has completed. This leaderboard reflects the final standings.							
#	Rank	Team Name	Notebook	Team Members	Score	Entries	Last
1	—	김주호			0.89182	30	4mo
2	—	송우철/나다			0.84613	10	4mo
3	—	홍재민			0.83908	34	5mo
4	—	hong			0.82954	8	4mo
5	—	99999			0.81883	27	4mo
6	—	이길호			0.80327	31	4mo
7	—	KKAng			0.80096	18	4mo
8	—	Hongheon Na			0.80079	13	4mo
9	—	Baseline_Code			0.87949	7	5mo
10	→ 4	xxx			0.87641	26	5mo
11	—	Bottom			0.87230	24	4mo
12	—	Sevan Yun			0.87200	38	4mo
13	→ 3	higher			0.86487	65	4mo
14	→ 1	한호호호호한호호호호			0.86041	30	4mo
15	→ 2	나호Moonheul			0.85968	49	4mo
16	→ 2	Lee won-jin			0.85953	25	4mo
17	→ 2	ja087			0.85826	5	5mo
18	→ 2	Yang Kyungju			0.85777	28	4mo
19	→ 2	KR_M mushroom			0.85219	8	5mo
20	→ 3	Jo++			0.84984	34	4mo
21	→ 1	고병진			0.84294	21	4mo
22	—	HeechulYang			0.84285	6	5mo
23	—	박기호			0.83663	3	4mo
24	—	JE1234			0.83311	30	5mo
25	—	홍하은홍시홍하			0.81564	8	4mo
26	—	MoonS			0.81168	14	5mo
27	—	로니리얼			0.79673	18	4mo
28	→ 5	U2019548			0.78350	11	4mo
29	→ 1	홍호호호호호호호			0.78115	27	4mo
30	→ 6	Kwon OhSang			0.78041	11	4mo
31	→ 1	U2019515			0.77987	22	4mo
32	→ 2	Yoon H			0.77987	41	4mo
33	→ 5	28me			0.77836	4	4mo
34	→ 5	apexch			0.77572	19	4mo
35	→ 4	HeeJung_Kang			0.77513	21	4mo
36	→ 1	Zenna			0.77484	5	5mo
37	→ 1	Eunyoung Yi			0.77425	11	5mo
38	→ 3	박지훈			0.77396	11	4mo
39	—	jeongheesa			0.77366	5	5mo
40	—	engyeEarth			0.77278	2	4mo
41	—	SM			0.77043	3	4mo
42	—	SIS			0.75488	5	5mo
43	→ 1	nae797979			0.75165	1	5mo
...

```
[2]: tr = pd.read_csv('train.csv')
     te = pd.read_csv('test.csv')
     tr.head()
```

```
t[2]:
```

	id	rating	row	chr_token	nn_token
0	R024795	1	1982년도에 미연영화가 제작이 된다는거 자체가 놀라운 거라고. 재미있다고 써부리는것...	1982년도에미연영화가제작이된다는거자...	년도영화제작거 자체가서 것 일찍 연속 장자...
1	R058332	1	현도별도기법을 적절히 써가며 현실과 환상의 대비를 확연하게 보여주는 영화	현도별도기법 적절히 써가며현실과환상의 대비를확...	현도별도기법현실 환상 대비 영화
2	R027236	1	...엑스맨을 아름답게 정리했다는 자체가 이미 뛰어난 영화라는 것을 증명하고 있다	...엑스맨을아름답게정리했다는자세가이미...	엑스맨 정리 자체 영화 것 증명
3	R050995	0	꼭 이런 쓰레기 영화가 다있나 인물관계도 정리안해줘줄거리도 없어 내용도 ㅋㅋ	꼭이런쓰레기영화가다있나인물관계도정리안해줄...	쓰레기 영화 인물 관계 정리안 줄거리 내용
4	R016317	0	육식공룡중에서 빌딩보다 더큰공룡이여잇노 ㅋㅋㅋㅋㅋㅋ	육식공룡중에서빌딩보다더큰공룡이여잇노ㅋㅋ...	육식공룡 중 빌딩 공룡 여잇노 ㅋㅋ

```
[3]: len(tr)
t[3]: 54510
```

```
[4]: len(te)
t[4]: 13626
```

```
[5]: cols = ['rating', 'nn_token']
     tr = tr[cols]
     cols = ['id', 'nn_token']
     te = te[cols]
     tr.head()
```

```
t[5]:
```

	rating	nn_token
0	1	년도 영화 제작 거 자체가 서 것 일찍 연속 장자 난자 천 리를리감 록 친구 여디...
1	1	현도 별도 기법 적절히 환상 대비 영화
2	1	엑스맨 정리 자체 영화 것 증명
3	0	쓰레기 영화 인물 관계 정리안 줄거리 내용
4	0	육식 공룡 중 빌딩 공룡 여잇노 ㅋㅋ

```
texts = list(pd.concat([tr.nn_token, te.nn_token]))
```

```
In [7]: from keras.preprocessing.text import Tokenizer
        from keras.preprocessing.sequence import pad_sequences

        Using TensorFlow backend.
```

```
In [8]: max_words = 10000
        tokenizer = Tokenizer(num_words=max_words)
        tokenizer.fit_on_texts(texts)

        seq = tokenizer.texts_to_sequences(texts)
        maxlen = max(len(x) for x in seq)

        word2idx = tokenizer.word_index
        word2idx[0] = 'PAD'
        idx2word = {j:i for i,j in word2idx.items()}

        seq = pad_sequences(seq, maxlen=maxlen)
        tr_seq = seq[:len(labels)]
        te_seq = seq[len(labels):]
```

```
In [9]: from sklearn.model_selection import train_test_split
        x_train, x_test, y_train, y_test = train_test_split(tr_seq, labels, test_size=0.2, stratify=labels)
```

```
In [10]: from keras.models import Sequential
         from keras.layers import Embedding, GRU, Dense

         embed_size = 100
         model = Sequential()
         model.add(Embedding(max_words, embed_size, input_length=maxlen))
         model.add(GRU(32))
         model.add(Dense(1, activation='sigmoid'))
         model.summary()
```

```
Model: "sequential_1"
```

Layer (type)	Output Shape	Param #
embedding_1 (Embedding)	(None, 91, 100)	1000000
gru_1 (GRU)	(None, 32)	12768
dense_1 (Dense)	(None, 1)	33

```
Total params: 1,012,801
Trainable params: 1,012,801
```

```
[ ] from transformers import pipeline
```

```
[ ] classifier = pipeline('sentiment-analysis')
```

Downloading: 100%  629/629 [00:00<00:00, 945B/s]

Downloading: 100%  232k/232k [00:01<00:00, 151kB/s]

Downloading: 100%  230/230 [00:01<00:00, 183B/s]

Downloading: 100%  268M/268M [00:05<00:00, 52.1MB/s]

```
[ ] classifier('더베피밤스터치 가맹점포 롯데리아 앞질러1위 우뚝')
```

```
[ ] [{"label": 'POSITIVE', 'score': 0.6384562848998474}]
```

```
[ ] classifier('맘스터치 대장군 초과 검출 이어 또 달 익은 파티, 맨손조리, 납킨 벌레 논란')
```

```
[ ] [{"label": 'POSITIVE', 'score': 0.7800015211105347}]
```

```
[ ] classifier('단독 지금도 맥도날드서 벌레, 이물질 줄줄줄반년새 21건 충격')
```

```
[ ] [{"label": 'POSITIVE', 'score': 0.7292629430361938}]
```

캐글(KAGGLE) Competition

2020-06 ~ 2020-12

국민대학교 경영대학원 AI빅데이터MBA

* 딥러닝 기말 Competition 및 대화형 AI, 데이터사이언스

실무 실습

* 감성분석 - 자연어 토큰화, LSTM, GRU, word2idx.시도

Jeon Eun Ha

THANK YOU!