



2021 빅콘테스트
2021 BIG CONTEST

제주도 음식물 쓰레기양 예측을 통한 배출량 감소 방안 도출

ECO제주 - 퓨처스리그

Team Name: 향긋한 쓰레기 유입량

최은재(팀장) aness0603@naver.com

김채현(팀원) 9808@naver.com

김하람(팀원) haramkim68@naver.com

최서윤(팀원) ab3492@naver.com





INDEX

ECO제주

1. 서론
2. 데이터
3. 탐색적 데이터 분석
4. 데이터 전처리
5. 모델링
6. 평가
7. 결론

#1

ECO제주

서론

1. 서론

HOME > 그린 리포트

온실가스 주범 음식물쓰레기...연간 885만톤 배출

이민선 기자 | 승인 2021.04.01 16:43

HOME > 행정 > 제주특별자치도

반복되는 제주 쓰레기 전쟁...늦장 대응만 계속 '시한폭탄'

김정호 기자 (newss@hanmail.net) | 승인 2021.08.29 16:05 | 댓글 14

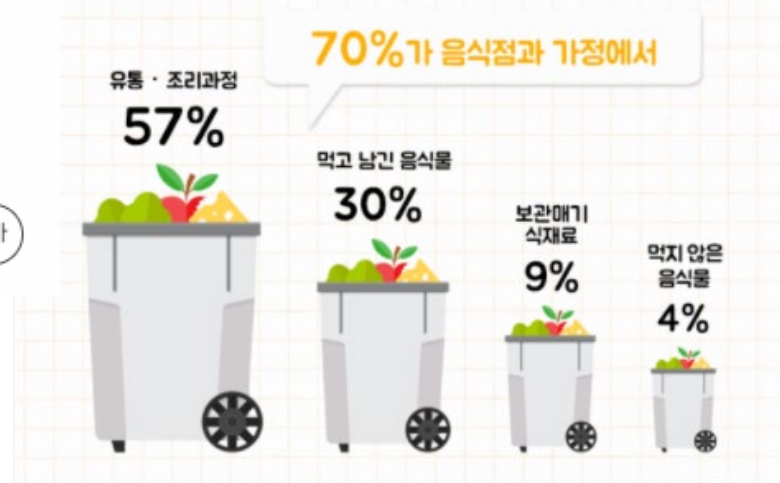


HOME > 오피니언 > 사설

음식물쓰레기, 처리가 아니라 줄여야 한다

뉴제주일보 | 승인 2020.09.22 19:44 | 댓글 0

쏟아지는 음식물쓰레기 1만 5천 톤!



#2

ECO계주



데이터

1) 문제정의

제주시지역 음식물 쓰레기 데이터(제주테크노파크)를 활용한

배출량 예측 모델 개발을 통하여

제주 읍면동별 배출량을 예측하고,

음식물 쓰레기 배출 요인에 따른

What are the factors?

배출량 감소 방안 도출

Plan

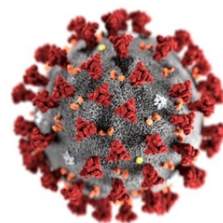
배출량 처리 방안 X
배출량 감소 방안 O

2. 데이터

2) 고려사항



Variable?

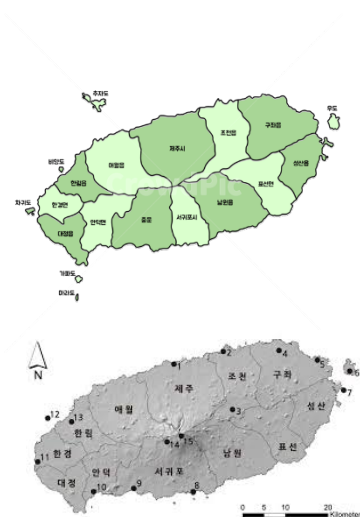


Covid 19

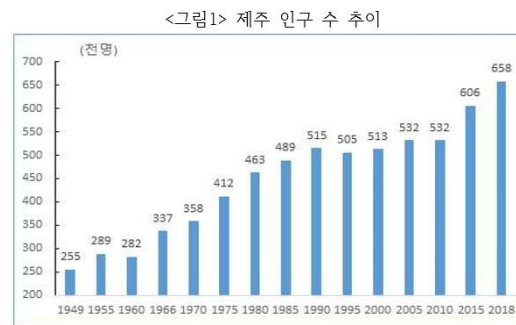


Holiday season

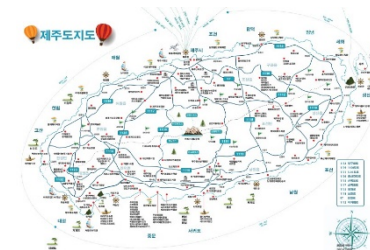
지역별?



지형별?



인구별? (남녀노소, 내외국인)



관광지별? (테마파크 or 주거집단)

2. 데이터

3) 데이터 개요

[대회 제공 데이터]

기간: 2018.01.01~2021.6.30



음식물 쓰레기



내국인
유동인구



장기체류
외국인
유동인구



단기체류
외국인
유동인구



거주인구



음식관련
카드소비

2. 데이터

3) 데이터 개요

[대회 제공 데이터]

기간: 2018.01.01~2021.6.30



음식물 쓰레기

No.	컬럼ID	컬럼명	사용여부
1	base_date	배출일자	X
2	city	시 구분	X
3	emd_cd	행정동 코드	X
4	emd_nm	행정동명	X
5	em_area_cd	배출거점지역 코드	X
6	em_cnt	배출건수	O
7	em_g	배출량	O
8	pay_amt	총 지불금액	O

2. 데이터

3) 데이터 개요

[대회 제공 데이터]

기간: 2018.01.01~2021.6.30



내국인
유동인구

No.	컬럼ID	컬럼명	사용여부
1	base_date	기준일자	X
2	time	시간	X
3	city	시 구분	X
4	emd_cd	행정동 코드	X
5	emd_nm	행정동명	X
6	sex	성별	X
7	age	연령대	X
8	resd_pop_cnt	거주인구	O
9	work_pop_cnt	근무인구	O
10	visit_pop_cnt	방문인구	O
11	resd	거주지역행정구역	X

2. 데이터

3) 데이터 개요

[대회 제공 데이터]

기간: 2018.01.01~2021.6.30



장기체류
외국인
유동인구

No.	컬럼ID	컬럼명	사용여부
1	base_date	기준일자	X
2	time	시간	X
3	nationality	외국인 국적코드	X
4	city	시 구분	X
5	emd_cd	행정동 코드	X
6	emd_nm	행정동명	X
9	resd_pop_cnt	거주인구	O
10	work_pop_cnt	근무인구	O
11	visit_pop_cnt	방문인구	O
12	resd	거주지역행정구역	X

2. 데이터

3) 데이터 개요

[대회 제공 데이터]

기간: 2018.01.01~2021.6.30



단기체류
외국인
유동인구

No.	컬럼ID	컬럼명	사용여부
1	base_date	기준일자	X
2	time	시간	X
3	nationality	외국인 국적코드	X
4	city	시 구분	X
5	emd_cd	행정동 코드	X
6	emd_nm	행정동명	X
7	visit_pop_cnt	방문인구	O

2. 데이터

3) 데이터 개요

[대회 제공 데이터]

기간: 2018.01.01~2021.6.30



거주인구

No.	컬럼ID	컬럼명	사용여부
1	base_year	기준년	X
2	base_month	기준월	X
3	city	시 구분	X
4	emd_cd	읍면동 코드	X
5	emd_nm	읍면동명	X
6	sex	거주자 성별	X
7	resid_reg_pop	주민등록 거주인구	O
8	foreign_pop	외국인 거주인구	O
9	total_pop	총 거주 인구	O

2. 데이터

3) 데이터 개요

[대회 제공 데이터]

기간: 2018.01.01~2021.6.30



음식관련
카드소비

No.	컬럼ID	컬럼명	사용여부
1	base_date	기준일자	X
2	city	시 구분	X
3	emd_cd	읍면동 코드	X
4	emd_nm	읍면동명	X
4	mct_cat_cd	음식 관련 업종 코드(대분류)	X
5	mct_cat_nm	음식 관련 업종명(대분류)	X
6	use_cnt	결제건수	O
7	use_amt	결제금액	O

2. 데이터

3) 데이터 개요

[외부 데이터]

기간: 2018.01.01~2021.6.30



국내선
항공노선별
운항통계



숙박업소
소비금액



제주도
입도객

2. 데이터

3) 데이터 개요

[외부 데이터]

기간: 2018.01.01~2021.6.30



국내선
항공노선별
운항통계

No.	컬럼ID	사용여부
1	기준일(월)	0
2	항공노선	0
3	도착출발구분	0
4	항공사	0
5	운항수(편)	0
6	이용자수(명)	0
7	화물량(톤)	0

2. 데이터

3) 데이터 개요

[외부 데이터]

기간: 2018.01.01~2021.6.30



숙박업소
소비금액

No.	컬럼ID	사용여부
1	일자	X
2	시도명	X
3	읍면동명	X
4	평균 기온	X
5	일강수량	X
6	최대 풍속	X
7	업종명	O
8	이용금액	O
9	데이터기준일자	X

2. 데이터

3) 데이터 개요

[외부 데이터]

기간: 2018.01.01~2021.6.30



제주도
입도객

No.	컬럼ID	사용여부
1	총계 당해년/전년 (월계/누계)	당해년 월계/당해년 증감율
2	총계 증감율 (월계/누계)	X
3	내국인 소계 당해년/전년 (월계/누계)	X
4	내국인 소계 증감율 (월계/누계)	X
5	내국인 개별여행 당해년/전년 (월계/누계)	X
6	내국인 개별여행 증감율 (월계/누계)	X
7	내국인 부분패키지 당해년/전년 (월계/누계)	X
8	내국인 부분패키지 증감율 (월계/누계)	X
9	내국인 패키지 당해년/전년 (월계/누계)	X
10	내국인 패키지 증감율 (월계/누계)	X
11	외국인 당해년/전년 (월계/누계)	X
12	외국인 증감율(월계/누계)	X

#3

ECO계주

탐색적
데이터
분석

3. 탐색적 데이터 분석

1) 결측치

[제공 데이터]

- 결측치가 없어 따로 보간하지 않음

[외부 데이터]

- 국내선 항공노선별 운항통계:

운항수(편)	9659
이용자수(명)	9659
화물량(톤)	9659
- 숙박업소 소비금액: 18/19/20년(21년 결측치 없음)

	[18]		[19]		[20]
평균 기온	6827	평균 기온	4979	평균 기온	302
일강수량	250	일강수량	1536	일강수량	135
최대 풍속	12303	최대 풍속	10320	최대 풍속	3268

* 사용하지 않는 컬럼이라 보간 진행하지 않음

3. 탐색적 데이터 분석

2) 시계열 데이터

[정상성 확인]

- 정상성: 시점에 관계없이 시계열 특성이 일정한 것
- 비정상 시계열 사용 시 예측 안정성 저하
- ADF 검정을 통한 판단 실시

```
#수식으로 정상시계열 판단하기
def adf_test(y):
    flag = True
    y = pd.to_numeric(y)
    df_test = adfuller(y, autolag='AIC')

    if df_test[1] <= 0.05:
        flag = True
        print(y.name, '는 정상시계열')
    else:
        flag = False
        print(y.name, '는 정상아님')

    return flag
```

3. 탐색적 데이터 분석

3) 음식물 쓰레기 배출량 확인

- 각 동별 배출량 확인
- 지역별 특성 고려



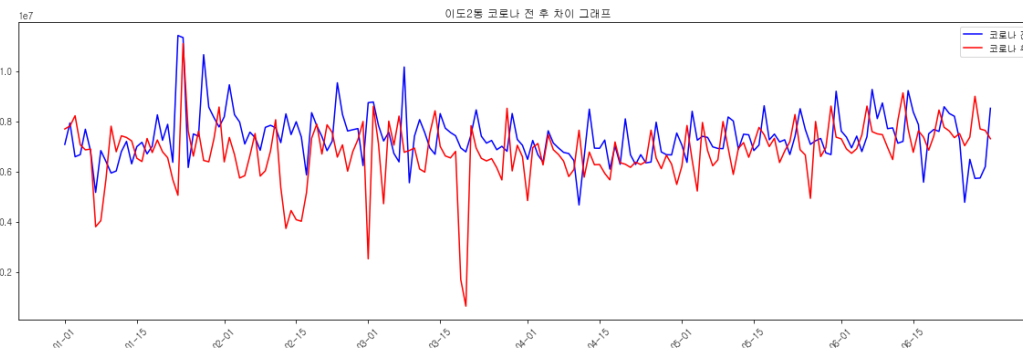
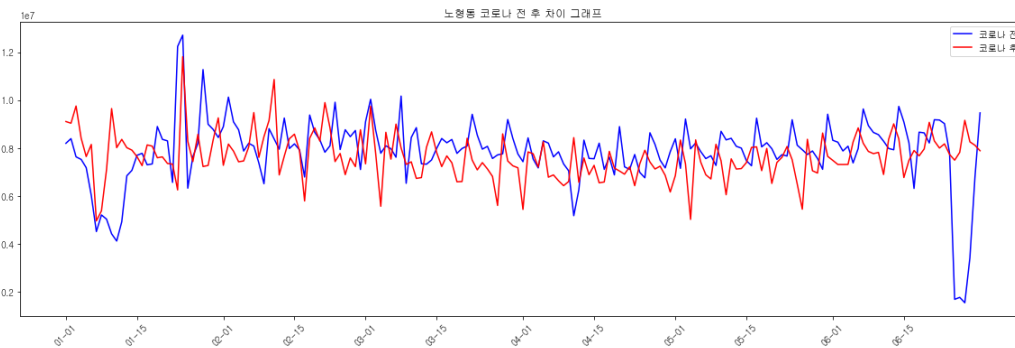
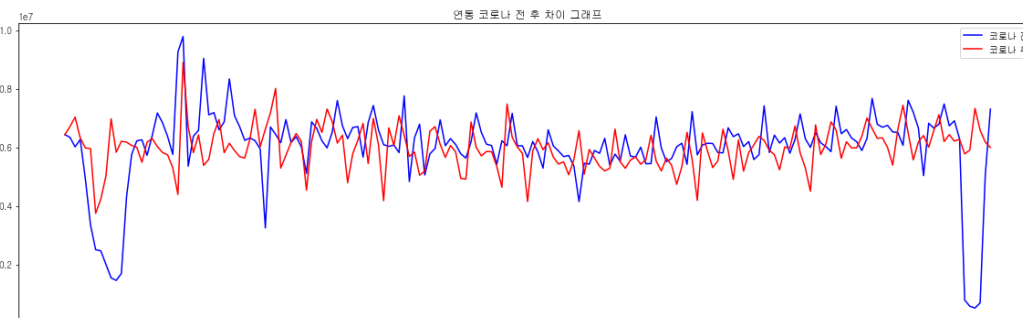
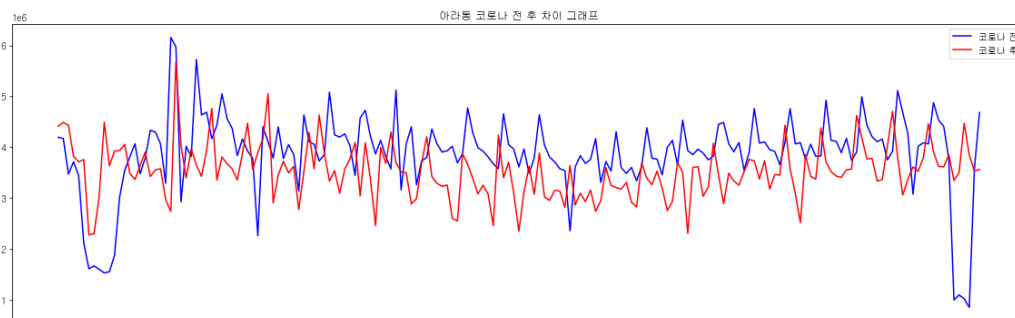
```
em_g_per_reg = {}
for i in em.groupby('emd_nm'):
    em_g_per_reg[i[0]] = sum(i[1]['em_g'])
em_g_per_reg = sorted(em_g_per_reg.items(), key=lambda x:x[1],reverse=True)
em_g_per_reg
```

```
(('노형동', 10598196902),
 ('미도2동', 9665869590),
 ('연동', 8282009221),
 ('일도2동', 6669949900),
 ('아라동', 5138943388),
 ('화북동', 4688385700),
 ('동홍동', 4377561131),
 ('삼양동', 4227416420),
 ('외도동', 4038596245),
 ('대정읍', 3402627328),
 ('용담2동', 2776403450),
 ('삼도1동', 2735122905),
 ('오라동', 2490460299),
 ('중문읍', 2365775726),
 ('남원읍', 2212342000),
 ('대천동', 2169291353),
 ('대륜동', 2088993653),
 ('미도1동', 1915102100),
 ('중앙동', 1898991600),
 ('성산읍', 1860060191),
 ('건입동', 1806250850),
 ('서홍동', 1768441400),
 ('애월읍', 1752144881),
 ('삼도2동', 1534443150),
 ('용담1동', 1528779299),
 ('천지동', 1310202150),
 ('한림읍', 1211911524),
 ('영천동', 1176418550),
 ('안덕면', 1160052747),
 ('송산동', 1087892600),
 ('표선면', 1045281485),
 ('조천읍', 977447096),
 ('정방동', 953359000),
 ('미호동', 853529300),
 ('일도1동', 795882200),
 ('도두동', 722142650),
 ('예래동', 629338266),
 ('봉개동', 624293449),
 ('효돈동', 587539550),
 ('구좌읍', 538528350),
 ('한경면', 349615500),
 ('알수없음', 151264969))
```

3. 탐색적 데이터 분석

4) 코로나 특성 고려

- 코로나 이전과 이후 데이터 분포 확인



#4

ECO제주

데이터
전처리

4. 데이터 전처리

1) 비정상 시계열 데이터 정상화

- 추세와 계절성 제거 필요
- 차분과 로그변환 실시 후,

ADF 재검정

resd_pop_cnt 는 정상아님
work_pop_cnt 는 정상아님
visit_pop_cnt 는 정상시계열

----- 전처리 후 -----

resd_pop_cnt 는 정상시계열
work_pop_cnt 는 정상시계열
visit_pop_cnt 는 정상시계열

----- 정상화 성공 -----

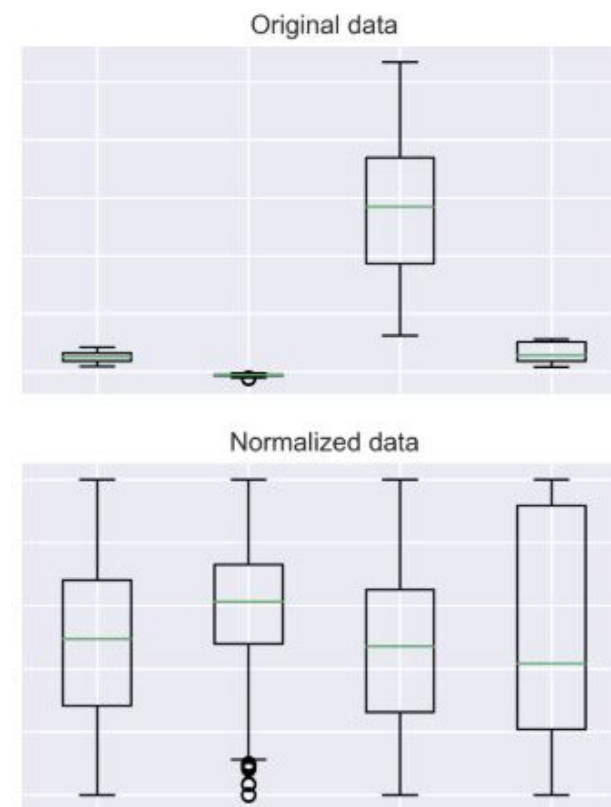
	time	emd_cd	resd_pop_cnt	work_pop_cnt	visit_pop_cnt
base_date					
2018-01-01	3024	12477529440	3669.0126	80.3557	2124.9315
2018-01-02	2771	12176866080	2885.6201	72.2265	1896.1328
2018-01-03	2430	11224765440	2859.4995	97.2012	1660.7993
2018-01-04	3065	12477529440	2613.4210	87.1216	2208.4086
2018-01-05	3771	15033168000	3149.9488	123.6984	2254.3613

4. 데이터 전처리

2) 데이터 정제

- 정규화: MINMAX Scaler 라이브러리 사용
- 데이터 스플릿: 8:2 비율로 Train/Test 분리

$$\frac{x - x_{min}}{x_{max} - x_{min}}$$



#5

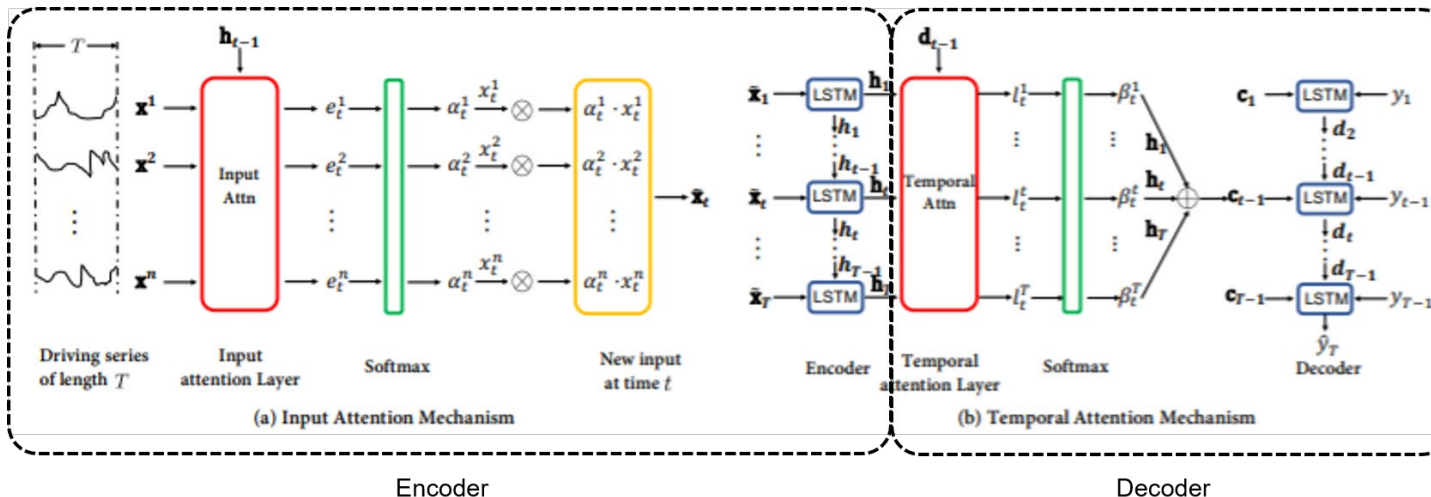
ECO제주

모델링

5. 모델링 - 딥러닝 후보

1) A Dual-Stage Attention-Based Recurrent Neural Network

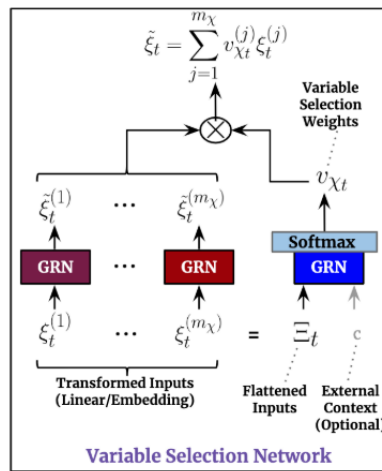
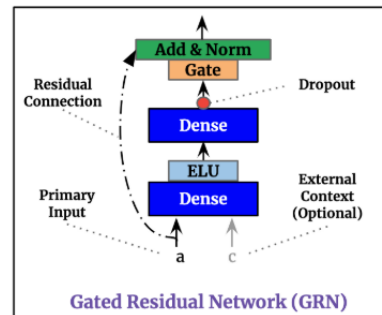
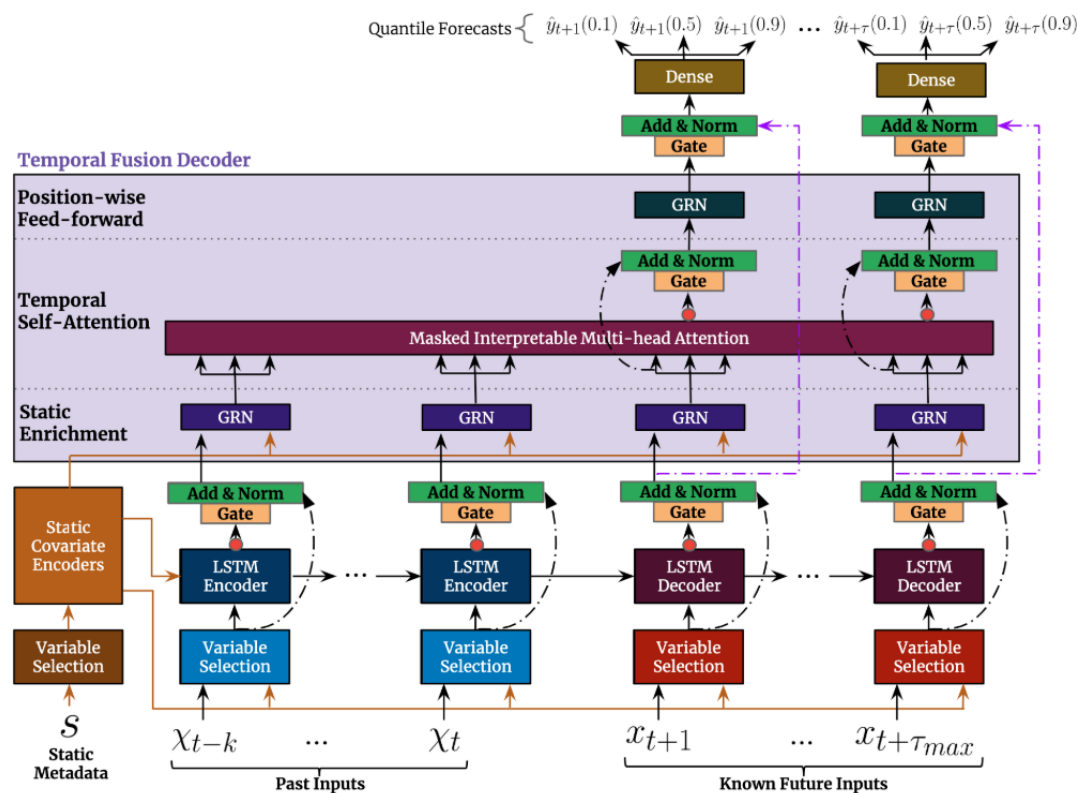
- ENCODER-DECODER 구조의 시계열 예측 모델
- 어텐션 기법을 이용해 FEATURE별, TEMPORAL별 중요도 산출 가능



5. 모델링 - 딥러닝 후보

2) Temporal Fusion Transformer

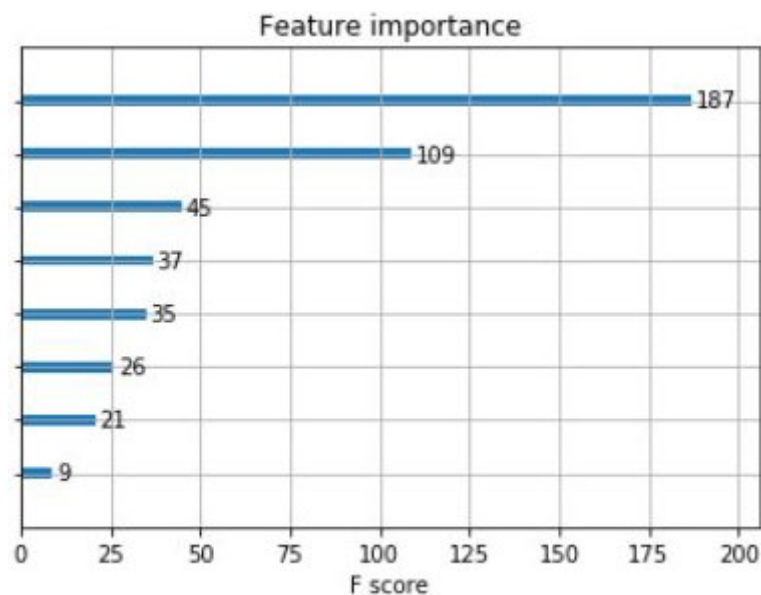
- 구글과 옥스포드 대학이 함께 만든 시계열 예측 모델
- 미래시점에 알 수 있는 데이터를 사용하여 예측 정확도 상승



5. 모델링 - 머신 러닝 후보

1) xgboost

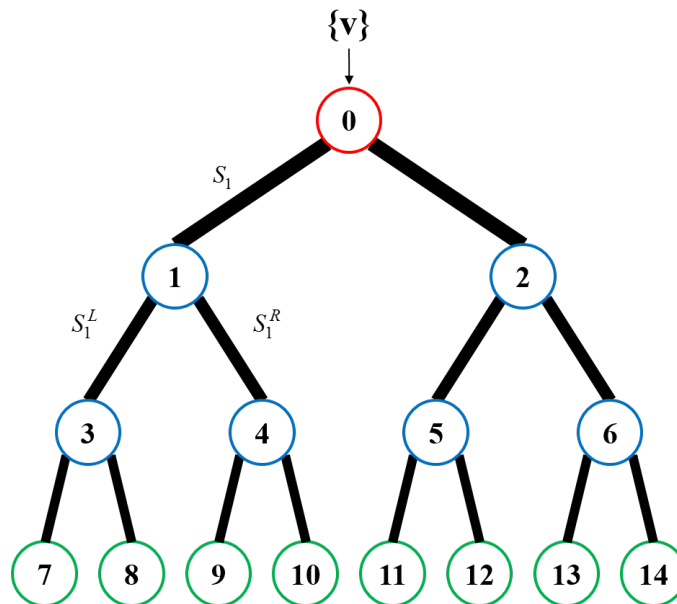
- 앙상블 모델로 결정 트리, 랜덤 포레스트 등의 기반을 가짐
- 중요 FEATURE 산출 가능



5. 모델링 - 머신 러닝 후보

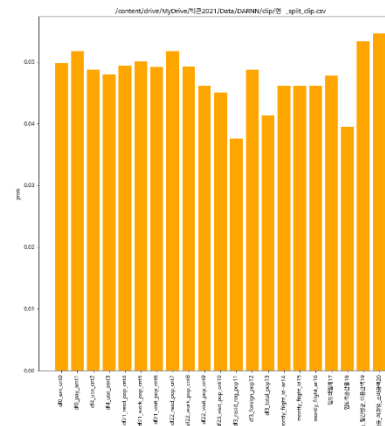
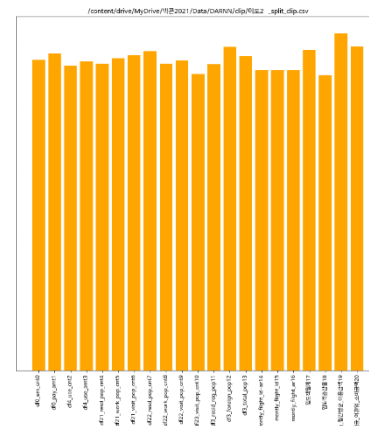
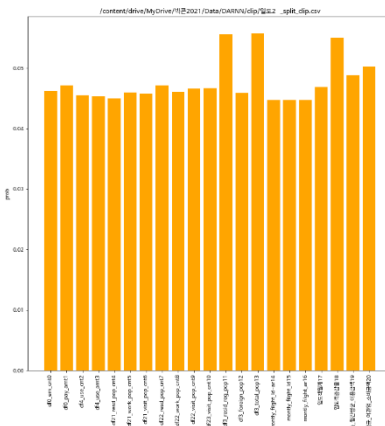
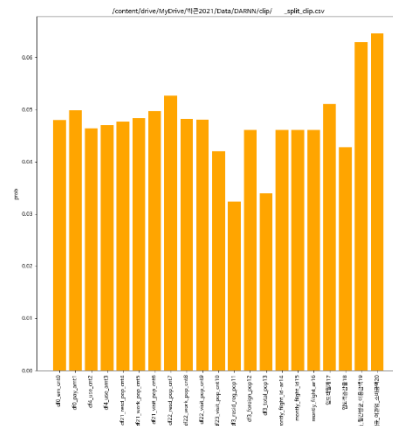
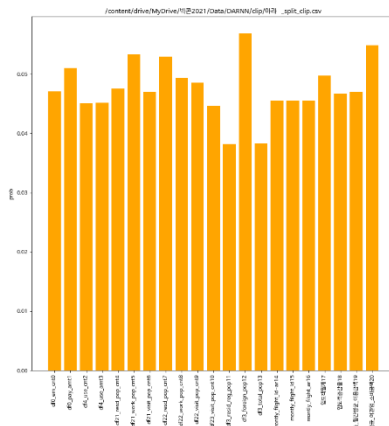
2) Random Forest

- 다수의 트리 모델에서 나온 결과를 취합하여 예측
- 회귀 모델로도 사용 가능



5. 모델링 - DA-RNN을 통한 변수 재설정

- 상위 5개 동의 ATTENTION WEIGHT 예시
- 해당 중요도를 참고하여 컬럼 선정에 사용



5. 모델링 - DA-RNN을 통한 변수 재설정

동 이름	사용 피쳐
노형동	em_cnt, pay_amt, use_cnt, use_amt, 국내_resid, 국내_work, 국내_visit, 외국단기_resid, 외국단기_work, 외국단기_visit, 거주_foreign, 월간항공편(도착-출발), 월간항공편_도착, 입도객월계, 모든업종_월간평균_이용금액, 월간평균_여관업_소비금액
이도2동	em_cnt, use_cnt, use_amt, 국내_resid, 국내_work, 국내_visit, 외국단기_resid, 외국단기_work, 외국단기_visit, 외국단기_visit, 거주_resid, 월간항공편(도착-출발), 월간항공편_도착, 월간항공편_출발, 입도객월계, 입도객증감율
연동	em_cnt, pay_amt, use_cnt, 국내_resid, 국내_work, 국내_visit, 외국단기_work, 외국단기_visit, 외국단기_visit, 거주_resid, 거주_foreign, 거주_total, 간항공편(도착-출발), 월간항공편_도착, 월간항공편_출발, 입도객증감율
일도2동	em_cnt, pay_amt, use_cnt, use_amt, 국내_visit, 외국단기_resid, 외국단기_work, 외국단기_visit, 외국단기_visit, 거주_resid, 거주_total_pop, 월간항공편_도착, 월간항공편_출발, 입도객월계, 입도객증감율, 모든업종_월간평균_이용금액
아라동	em_cnt, pay_amt, use_cnt, use_amt, 국내_resid, 외국단기_resid, 외국단기_work, 외국단기_visit, 거주_resid, 거주_total, 월간항공편(도착-출발), 월간항공편_도착, 월간항공편_출발, 입도객월계, 입도객증감율, 모든업종_월간평균_이용금액

5. 모델링 - 최종 선정 컬럼

TFT	제공데이터 + 월간항공편(도착-출발), 월간항공편_도착, 월간항공편_출발, 입도객월계, 입도객증감율
DARNN	이전 슬라이드
머신러닝	제공데이터 + 외부데이터

제공데이터 : em_cnt, em_g, pay_amt, use_cnt, use_amt, 국내_resid, 국내_work, 국내_visit, 외국단기_resid, 외국단기_work, 외국단기_visit, 외국단기_resid, 외국단기_work, 외국단기_visit, 거주_resid, 거주_foreign, 거주_total

외부데이터 : 월간항공편(도착-출발), 월간항공편_도착, 월간항공편_출발, 입도객월계, 입도객증감율, 모든업종_월간평균_이용금액, 월간평균_여관업_소비금액

선정방법

- DARNN : 피쳐중요도 하위 5개 제외
- TFT : 모든 동에 존재하는 데이터만 포함(메타데이터)
- 머신러닝 : 제공데이터 + 동별로 존재하는 외부데이터

5. 모델링 - 최종 선정 모델 및 예측

- 예측 그래프를 보며 성능을 테스트한 결과 최종적으로 3가지 모델을 사용하기로 결정
- 최종 선정 모델
 - DA-RNN
 - XGBoost
 - Random Forest
- 예측 방법
 - 세 모델의 예측 값을 더하여 평균 값으로 산출
$$\frac{(DA-RNN\ result)+(XGBoost\ result)+(Random\ Forest\ result)}{3}$$

3

#6

ECO계주

평가

6. 평가

1) MSE

2) MAE

- 두가지 평가지표를 통해 예측값과 실제값이 얼마나 차이 나는지 판단

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2$$

$$MAPE = \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{\hat{y}_i} \right|$$

6. 평가

[DA-RNN]

mse: 0.0048060156

mae: 0.055508606

[Random Forest]

mse= 0.02348

mae= 0.12535

[XGBoost]

mse= 0.019206

mae= 0.11233

최종 예측 결과는

3가지의 모델의 결과를 조합

#7

ECO제주

결론

수도권 지역의 음식물쓰레기 발생 모니터링 결과 비교

전태완 · ○김용준 · 정미정 · 신선경 · 박종은 · 정용우

국립환경과학원 환경자원연구부 자원순환연구과

행정통계는 폐기물을 수거하여 처리되는 단계에서
파악되는 처리량 통계만 파악되고
발생된 폐기물의 효과적인 처리방안을 결정하기 위해서는
단순한 폐기물 발생량에 대한 조사뿐 아니라
발생원 단계의 폐기물 종류에 따른 발생량의 조사가 중요.

즉 가정이라는 기초 발생원에서 직접 배출되는
음식물쓰레기의 양, 종류, 보관형태, 처리기한, 배출형태 등
기초자료 확보와 성상파악이 감량화 성과의 구체적 자료 도출,
체계적인 수거체계, 음식물쓰레기 저감 촉진 방안 마련을 위해
가정에서의 음식물쓰레기 배출 실태파악이 필요.

※ 단순히 통계로 잡히는 음식물 쓰레기 뿐만 아니라
발생원 단계의 폐기물에서도 조사가 필요함.

음식물쓰레기 재활용 정책의 분석과 과제*

최창원**·정윤수***·이진원****·김월중*****

음식물류 폐기물 재활용현황 분석을 통한 사업화 연구

박용수(주식회사 청호산업)*
설병문(경남과학기술대학교 창업대학원)**

음식물류폐기물 종합관리체계 구축을 위한 배출부터 이용까지의 문제점 도출 및 개선사항 검토

○서요평¹·김영신²·조준연²·배재근²

서울과학기술대학교 일반대학원 환경공학과¹,

서울과학기술대학교 에너지환경대학원 에너지환경공학과²

음식물쓰레기에 이물질이 많고, 자격이 안 된 재활용 처리 업체로 인해 음식물쓰레기를 재활용한 제품(사료, 퇴비)의 품질이 낮아 사용을 기피 한다는 것.

아울러, 유통체계가 체계적이지 못하고, 정부의 관심부족으로 일부 품질 좋은 제품의 경우에도 필요한 농가에 적시에 제공되지 못하고 있는 것.

※ 음식물쓰레기 분리배출기준을 명확히 하고
비닐봉투를 금지하며, 부실업체의 퇴출제도를 확립.

농산물 도매시장의 음식물류 폐기물은 원상반출 및 탈수 보다는 탈수 후 건조가 처리능력을 향상 시킬 수 있는 것으로 판단.

그리고 탈수 후 건조된 음식물류 폐기물은 양질의 퇴비화 원료를 활용할 수 있기 때문에 농산물 도매시장의 음식물류 폐기물은 탈수 후 건조하여 퇴비화 방법으로 재활용사업화를 추진.

ECO제주



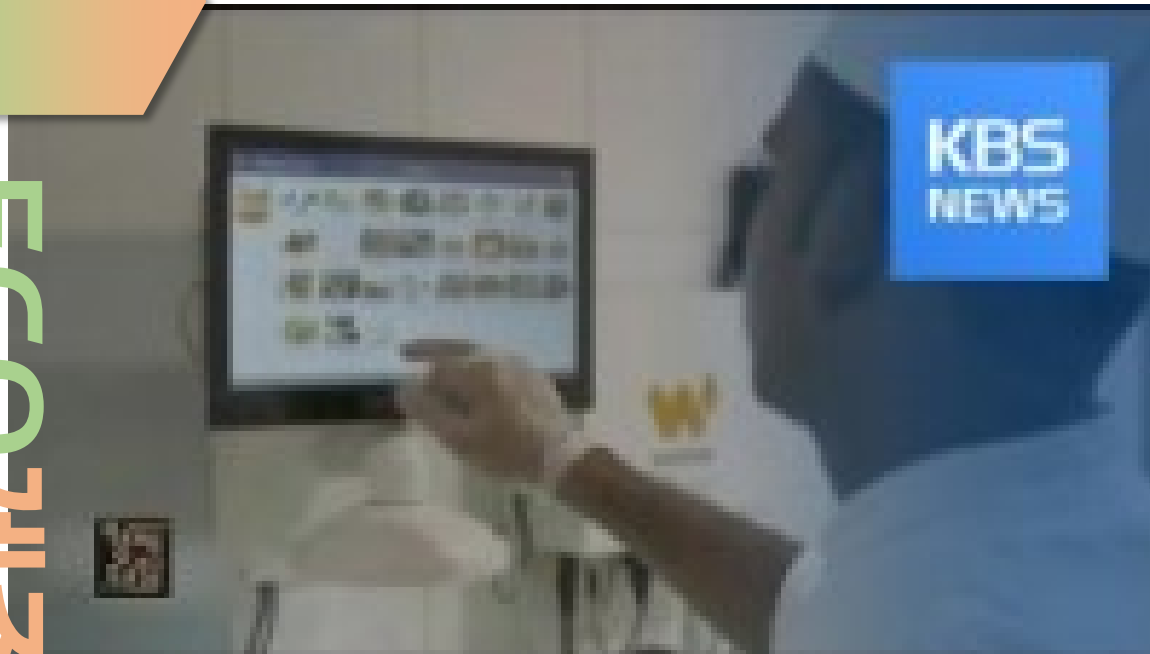
식판에 줄을 그어 음식량을 조절.

결과: 잔반 70% 줄어듦, 기존 식판 무게도 3분의 1로 축소



80%는 미생물로 분해

결과: 음식물 쓰레기 양을 20%로 줄임, 주택가에 효과적



유통기한 임박상품 주요 판매처

라스트오더	모바일 앱으로 주변 매장 할인정보
떠리몰	신선·가공식품, 건강기능식품 등 판매
달달몰	과자 사탕 초콜릿 등 당류 제품 판매
이유몰	식품, 과자, 건강기능식품 판매
다모고	음식점들의 마감할인 정보 제공

유통기한 임박 상품을 판매하는 플랫폼 제시

- 유통기한이나 마감이 임박한 상품들을 스마트폰 앱을 통해서 할인 판매.
- 결과: 매장은 매출 증가, 소비자는 저렴하게 구입.



RFID기반 음식물쓰레기 관리체계 구축·운영

RFID태그를 인식 후 음식물쓰레기를 배출하면 배출자, 배출량(무게), 시간 정보가 중앙시스템으로 전송되어 수수료를 부과하고 발생량과 감량 등 통계 데이터를 관리.

中 먹방 하면 벌금 1700만원... 음식낭비금지법 발의

✎ 김은지 기자 | Ⓜ 입력 2020.12.23 16:52 | 💬 댓글 6

ECO저널

역시 중국.. 금지법으로 철폐.

다만 여기서 살펴볼 점!!

“이와 같은 이유로 초안된 음식 낭비 금지법안이 통과되면 중국에서 인기를 끄는 콘텐츠인 먹방(먹는 방송)도 제한되게 된다. 만약 먹방 콘텐츠를 방송하면 시정명령과 함께 1만~10만 위안(170만~ 1천700만 원)의 벌금까지 부과된다고 전했다.”

※ 메스컴에 오르는 먹방, 푸드파이트 이런 매체적 영향들이 시청자로 하여금 심리적인 영향을 주어 폭식이나 과식을 유도할 수도 있다고 볼 수 있음.



효율적인 음식물 쓰레기 처리 방안 - 환경부

요인

1. 음식물에 대한 심미적 기준
2. 배송 과정에서의 협력 부족
3. 개인들의 실천 부족
4. 일관되고 지속적인 정책의 부재

방안

1. 음식물에 대한 지나친 심미적 기준 완화
2. 배송 시스템 혁신
3. 출장 식품업 혁신
4. 음식물 기부 장려

7. 최종 결론

1. 정부

- 개인과 기업이 음식물 쓰레기를 줄일 수 있도록 **사전 교육**
- 남은 음식물이 버려지지 않도록 기부 또는 거래할 수 있는 **플랫폼 구축**
- 기업(가게)이 음식물 쓰레기 배출량을 줄일 경우 그에 상응하는 **인센티브** 지급 방안
- 음식물 쓰레기의 부피자체를 줄일 수 있도록 **수분을 제거**할 수 있는 쓰레기통 개발
- 먹방 관련 세부 규정 필요(일반 소비자에게 경각심을 줄 수 있는 문구를 영상에 포함하는 등)

2. 개인

- 음식물을 **적정량만 소비**할 것
- 사용하는 식재료를 미리 계산하여 **계획적 구매**

3. 기업(또는 가게)

- **과도한 밑반찬** 제공을 자제하기
- 음식양에 따른 **메뉴 세분화**