

배달과 보험의 변화

헬라 클레스

멀티캠퍼스

융복합 프로젝트형: 융합 프로젝트과정

Team 헬라클레스

박성준 이동규 이재환 이주호 홍연하



목차

01.

PJT 개요

- 환경분석: 기획 배경
- 사용자/엔드 유저 분석
- 아이템 소개
- 목적 및 기대효과

02.

프로젝트 수행 내역

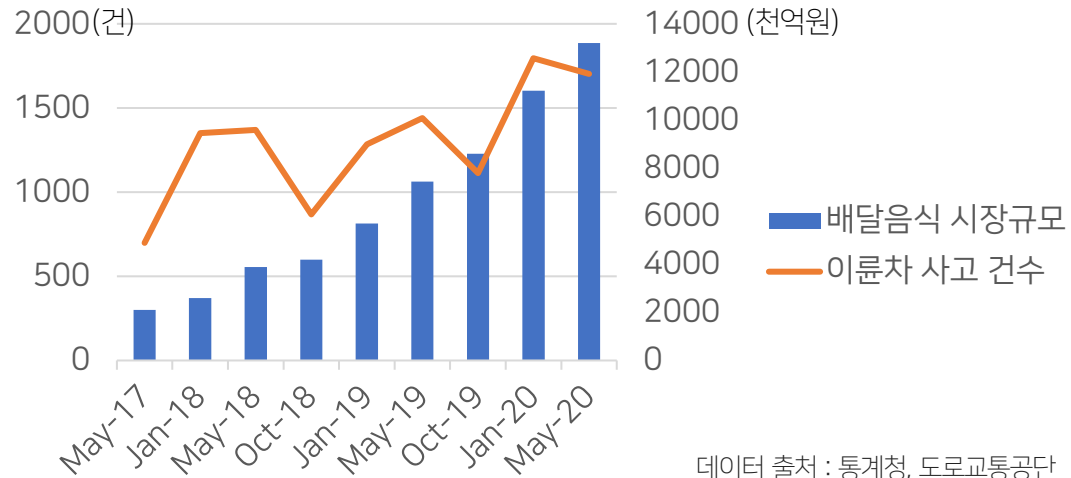
- 팀 구성 및 역할분배
- 시연 영상
- 수행 절차 및 방법

03.

For next step!

- 보완 되어야할 사항들

안전, 보험 그리고 수익 사이에서 사각지대에 놓여있는 있는 배달원들



1. 시장상황

코로나19 영향으로 음식 배달거래액 2019년 대비 83% 증가

→ 배달원들에 대한 수요증가

→ 배민 커넥트 등 쉽게 단기 목적으로 배달대행가능



<한정애 의원, 배달사고에도 중대재해 조사토록 규칙 개정해야>
출처 : 서울로컬뉴스(<http://www.slnews.co.kr>)

2. 배달원

- 늘어나는 배달원, 늘어나는 교통사고.
- 배달원 관련 보험과 정책들이 논의되고 있음.
- 다양한 교통 수단으로 배달할 수 있게됨.

사회적 이슈가 되고 있는 배달원 안전문제



라이더/안전에 대한 뉴스기사 162
건에 대한 내용 워드 클라우드 분석

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift
0	(배달)	(라이더)	0.872727	0.172727	0.172727	0.197917	1.145833
1	(라이더)	(배달)	0.172727	0.872727	0.172727	1.000000	1.145833
2	(라이더)	(안전)	0.172727	0.554545	0.109091	0.631579	1.138913
3	(안전)	(라이더)	0.554545	0.172727	0.109091	0.196721	1.138913
4	(배달)	(민족)	0.872727	0.109091	0.109091	0.125000	1.145833
5	(민족)	(배달)	0.109091	0.872727	0.109091	1.000000	1.145833
6	(사고)	(배달)	0.136364	0.872727	0.136364	1.000000	1.145833
7	(배달)	(사고)	0.872727	0.136364	0.136364	0.156250	1.145833
8	(배달)	(안전)	0.872727	0.554545	0.472727	0.541667	0.976776
9	(안전)	(배달)	0.554545	0.872727	0.472727	0.852459	0.976776

라이더에 대한 뉴스기사 185건에 연관분석

더 안전하게 운전할 수는 없는 걸까?

비오는 날 넘어져 주저앉은 배달원

직접 체험해봤습니다.



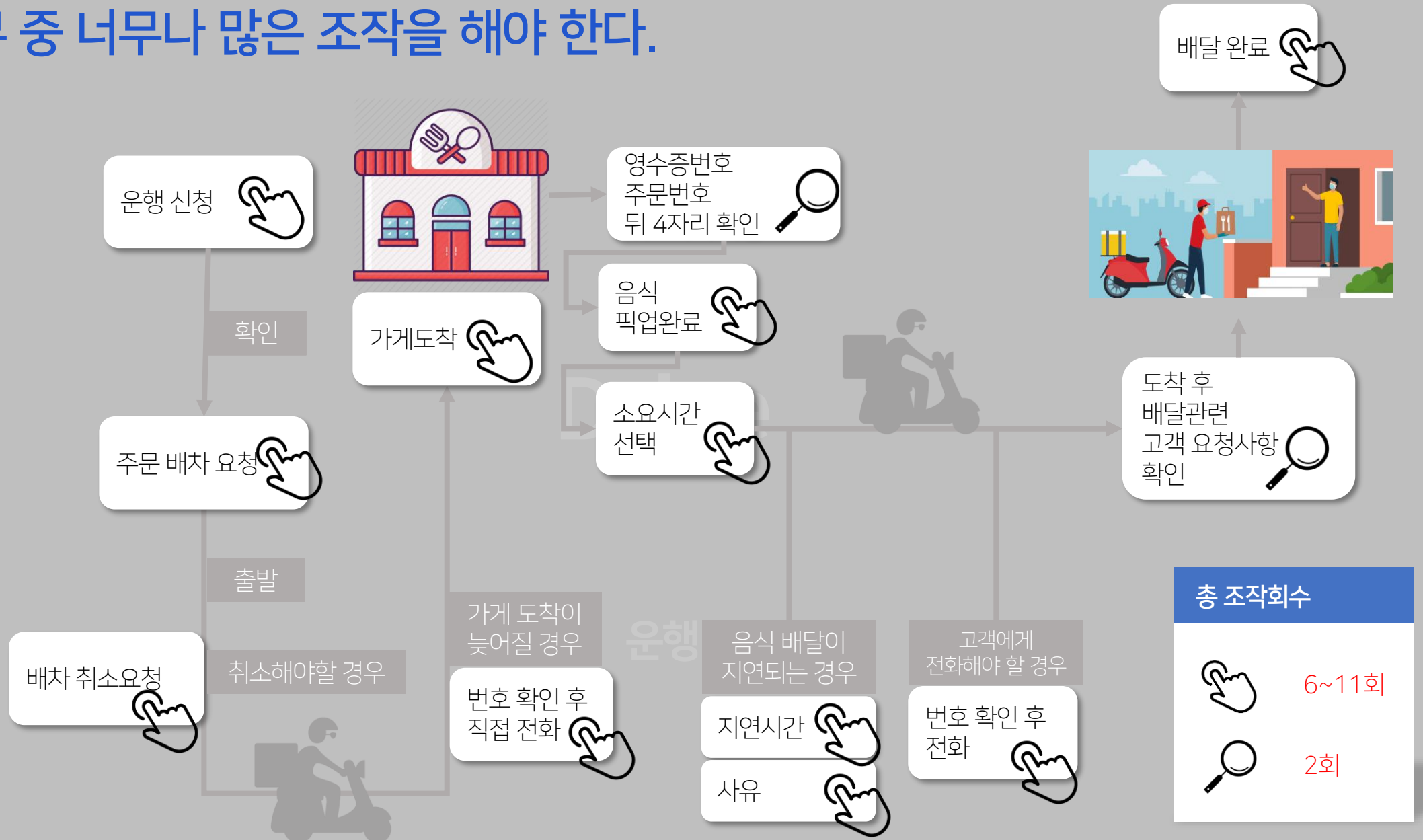
Drive

운행

Cost

비용

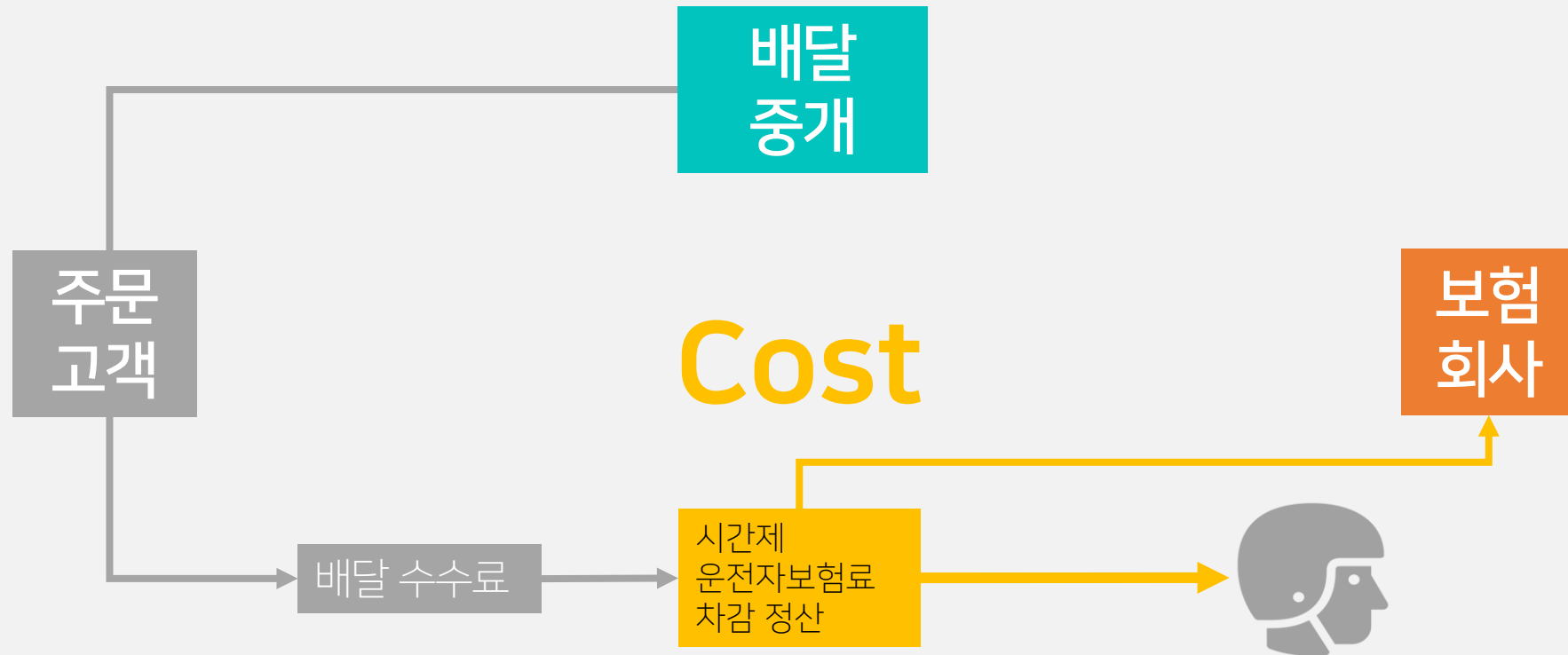
업무 중 너무나 많은 조작을 해야 한다.



비용

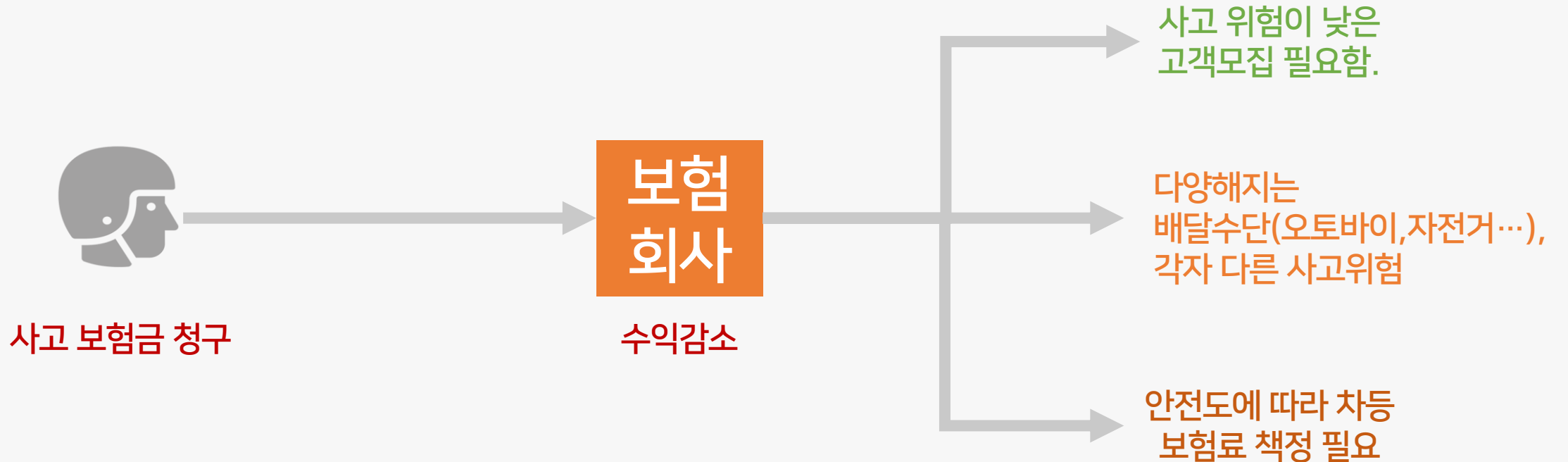
Cost

배달 수수료에서 시간제 운전자 보험료를 차감하여 수익정산



보험회사 입장에선 사고가 발생할 수록 손해

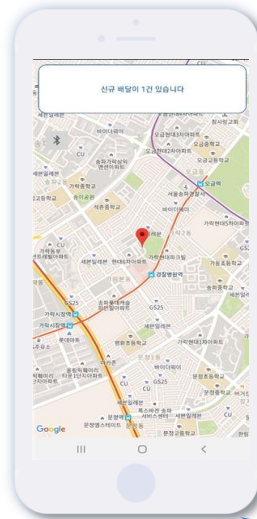
1.2 사용자분석: 보험회사



음성인식으로 안전한 배달업무 개선과

보험용 운행데이터를 수집하는 서비스

음성인식 챗봇



"음식 픽업 완료"

"네 ** 주문 픽업완료
처리했습니다."

"안전운행 하세요."

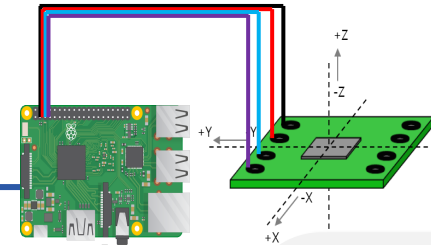


이상치 감지

가속도
자이로 센서



DB: 배달원
운행데이터



- 급정지, 급커브, 급 U턴 등으로 의심되는 센서의 이상치 데이터가 발견되었을 때
 1. 배달원에게 알려줌.
 2. 서버에 데이터를 축적함
- 데이터를 다양하게 활용 할 수 있도록 서버에 전송합니다.

시연 영상

[로그인 및 운행시작]
블루투스 연결 및 픽업지 확인



아이디
test

비밀번호

회원가입

모니터링 페이지

영상

기능1. 음성인식으로 안전한 배달업무 개선

음성인식 챗봇



"음식 픽업 완료"

"네 ** 주문 픽업완료
처리했습니다."

"안전운행 하세요."

아이템 기능

배달 업무 중 발생하는 앱 조작을 음성인식을 통해 처리하고,
안전운전을 유도하는 기능

1. 배달 중 앱 조작을 터치에서 음성조작으로 변경해 화면을
보지않아도 일할 수 있게
2. 센서에서 진동 감지, 과하게 흔들릴 때 라이더에게 알려줘
안전운전과 온전한 음식배달을 유도함.

기능2. 배달원의 안전운전을 파악할 수 있는 데이터 수집 IOT

	id	userID	shopName	comAddress	destination	deliveryTime	alertCount	comLatitude	comLongitude	desLatitude	desLongitude	distance	assignDate	status	age	rank
0	1	user1	초록마을	서울 강남	서울시 강	20.5	6	37.48894	127.0682	37.50935	127.0389	3.445286	2020-11-09 18:00:00	0	27	100
1	2	user2	신현대상	서울 강남	서울시 강	22.8	4	37.47916	127.0497	37.50935	127.0389	3.489975	2020-11-09 20:00:00	0	43	100
2	3	user3	포이현대	서울 강남	서울시 강	18.7	4	37.47916	127.0497	37.50935	127.0389	3.489975	2020-11-09 22:00:00	0	25	100
3	4	user4	포이현대	서울 강남	서울시 강	19.2	9	37.47916	127.0497	37.50935	127.0389	3.489975	2020-11-10 00:00:00	0	27	95
4	5	user5	주식회사	서울 강남	서울시 강	20.5	5	37.51576	127.0326	37.50935	127.0389	0.899932	2020-11-10 02:00:00	0	58	85
5	6	user6	빙달(논)	서울 강남	서울시 강	20.2	6	37.51842	127.038	37.50935	127.0389	1.010807	2020-11-10 04:00:00	0	32	85
6	7	user7	세븐브릭	서울 강남	서울시 강	28.9	0	37.51635	127.0379	37.50935	127.0389	0.783285	2020-11-10 06:00:00	0	22	110
7	8	user8	계화기식	서울 강남	서울시 강	26.4	1	37.5168	127.0381	37.50935	127.0389	0.830566	2020-11-10 08:00:00	0	46	100
8	9	user9	네네치킨	서울 강남	서울시 강	25.5	2	37.51012	127.0356	37.50935	127.0389	1.173198	2020-11-10 10:00:00	0	58	95

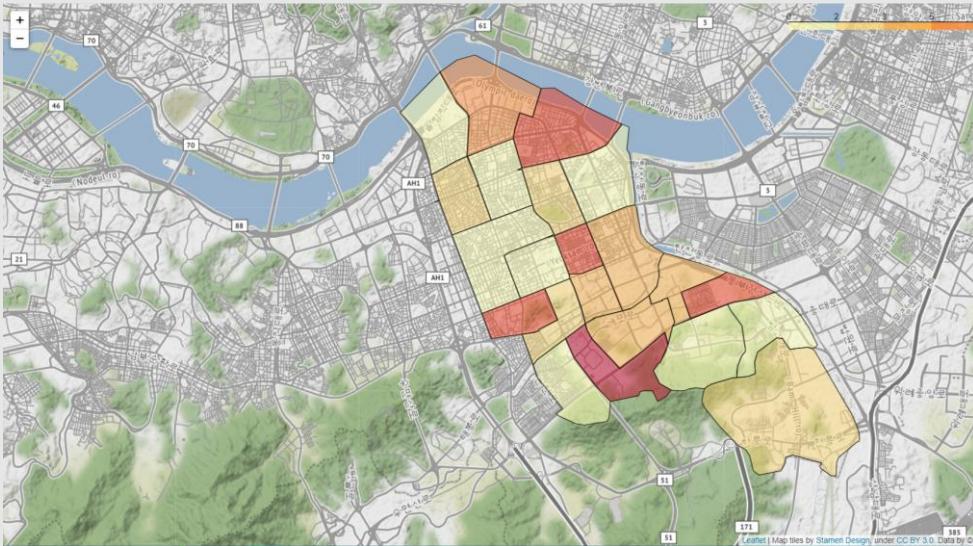
배달원 운행 데이터

보험회사가 사용할 수 있는 배달원의 운전데이터를 수집하고 제공

1. 주요 수집 데이터: 배달 출발지/ 도착지/ 배달시간/ 이상치 발견회수/ 운행거리 등.
2. 배달원의 안전운행지수를 이용하여 보험상품 설계에 도움 : 안전지수에 따른 차등 보험료 책정 등
3. 배달원이 안전하게 운전할 수 있게 하고 사고율을 줄여서, 보험금 지급 확률을 낮춤

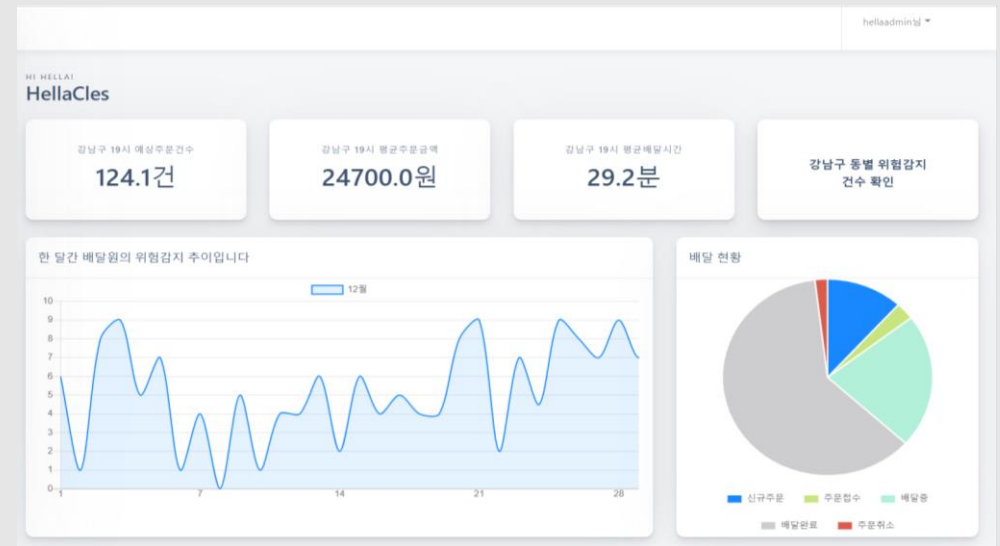
모니터링 페이지

시간, 지역별 이상치 발생



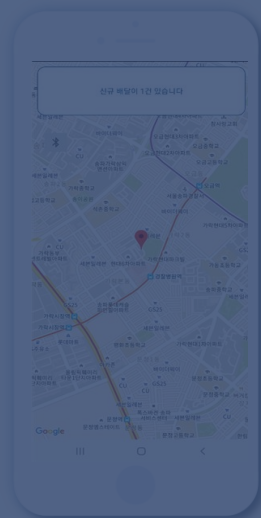
- 시간, 지역별 위험신호 발생정도를 확인
- 라이더의 운전지역/ 시간에 따라 사고 위험도를 평가할 수 있음.

배달원 안전 운행 개선정도 평가



- 배달원들의 위험 운전 정도를 모니터링해 '헬라'가 배달원들의 안전운전에 얼마나 영향을 주는지 확인할 수 있습니다.

음성인식 챗봇



"음식 픽업 완료"

"네 ** 주문 픽업완료
처리했습니다."

"안전운행 하세요."

PJT 수행내역

팀 역할 분담

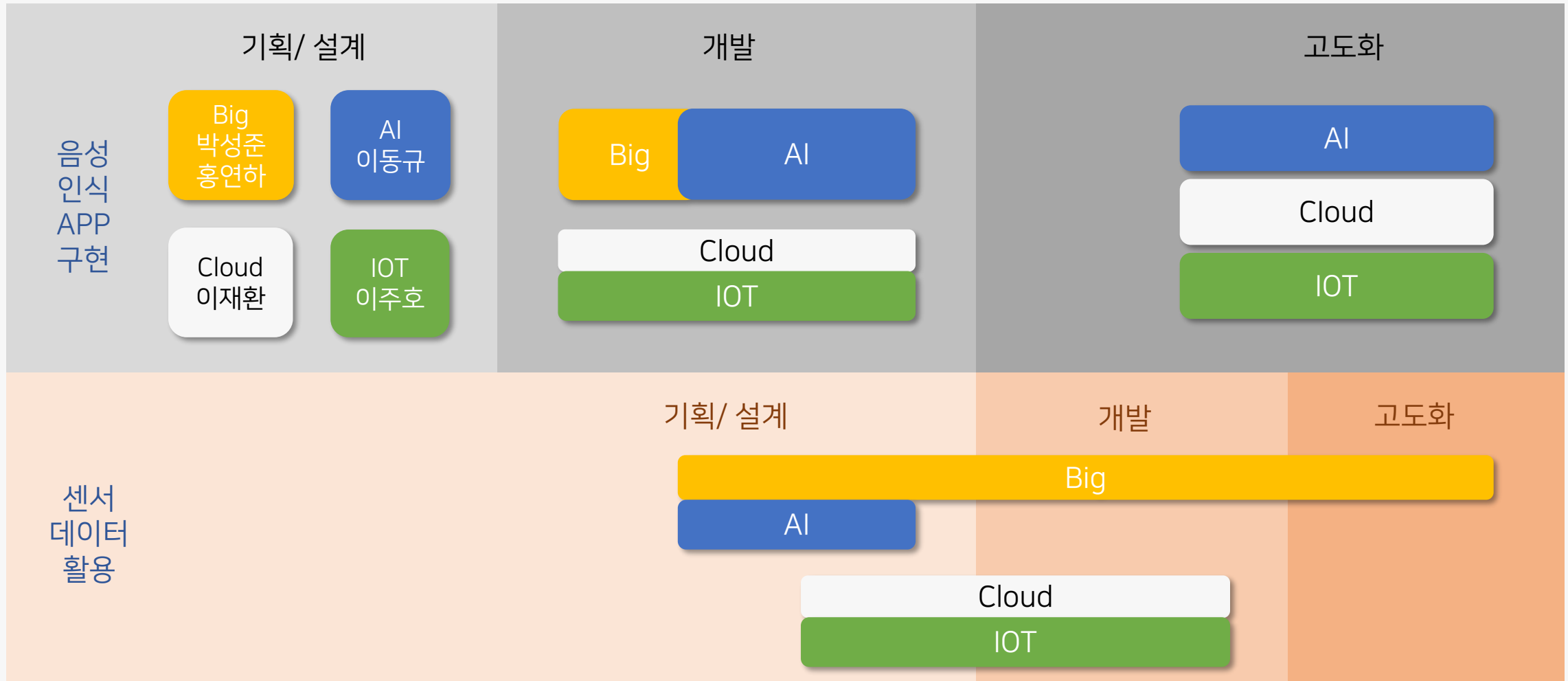
이상치 감지

가속도
자이로 센서

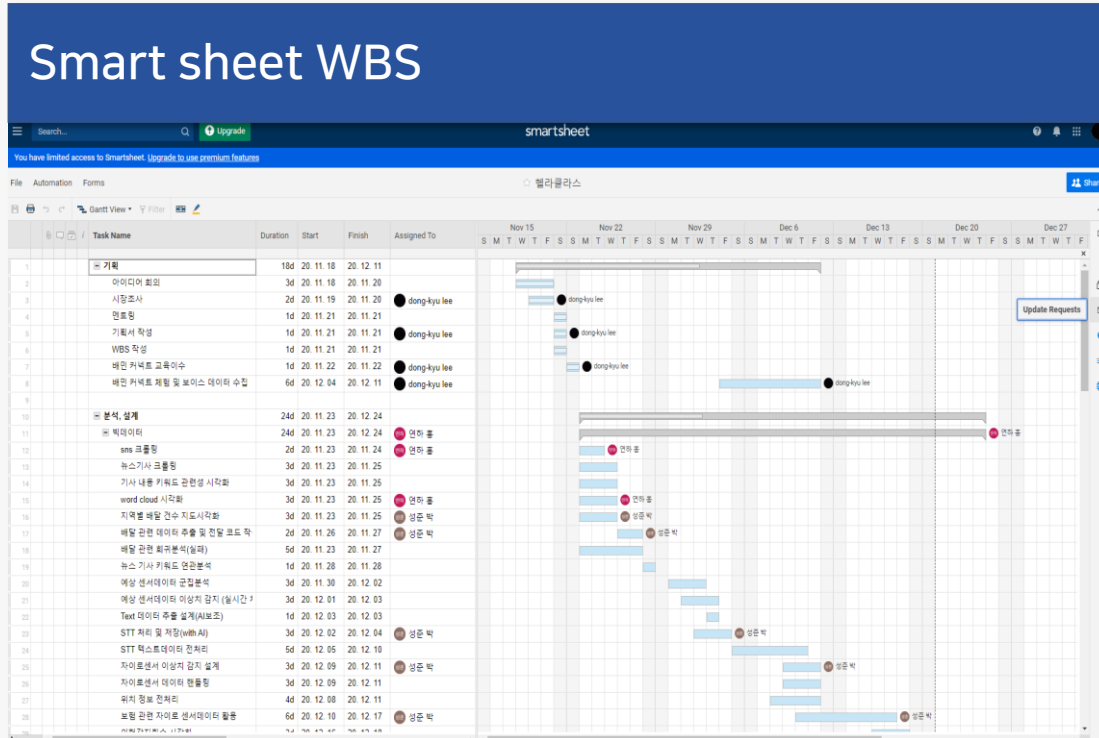


DB: 배달원
운행데이터

PJT 역할 분담.

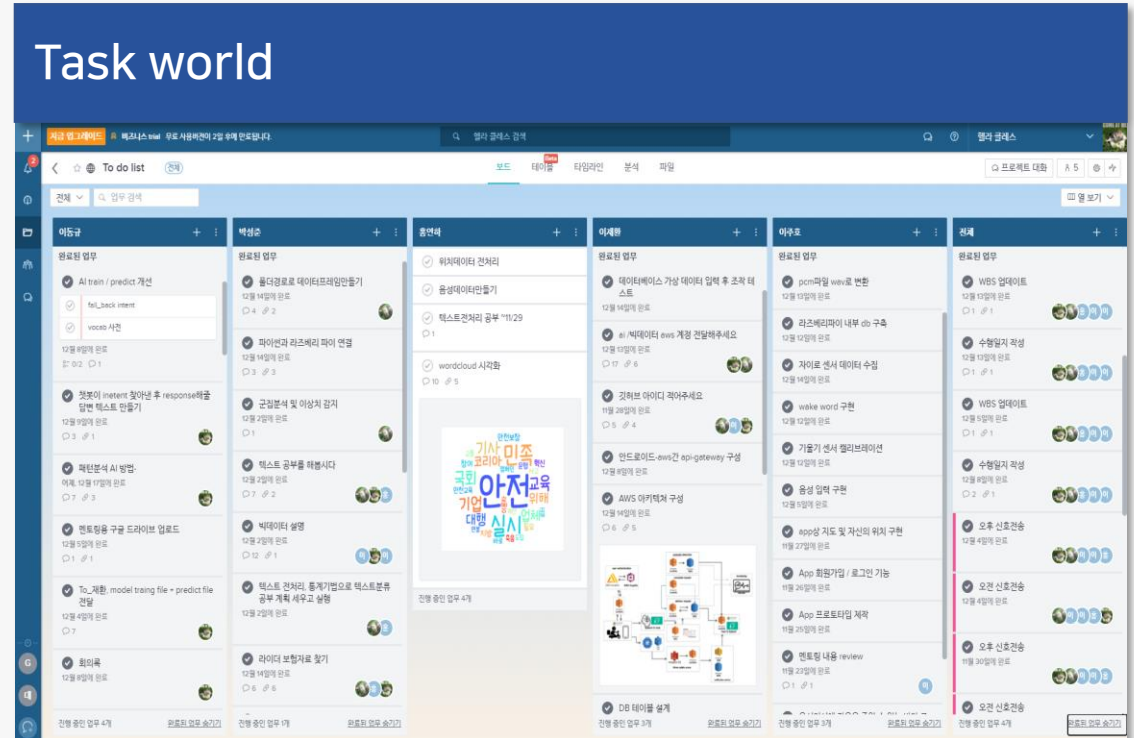


협업 방식



Planning

- WBS로 전체/ 주차 별 계획



Communication

- 칸반보드 협업 툴 이용, 원활한 업무 분배와 커뮤니케이션 도모

빅데이터 - 홍연하

뉴스기사 크롤링



라이더 안전 키워드로 뉴스기사 제목 크롤링

0	"라이더 안전 위한 법 만들어야"...국회 찾은 배달 기사들
1	"배달하다 과당! 안돼요" 라이더 안전 챙기는 기업들 - 아주경제
2	배달의민족, 서울경찰청과 라이더 안전 교육 실시 - 지디넷코리아
3	물 만난 배달대행업체?...라이더 안전은 '빨간불' - 노컷뉴스
4	안전한 배달 위해선 '라이더안전보장법' 필요 - 참여와혁신
...	...
157	안전한 배달 위해선 '라이더안전보장법' 필요 - 참여와혁신
158	배달의민족, 서울지방경찰청과 함께 라이더 안전교육 실시

```
from konlpy.tag import Okt
from collections import Counter

news_corpus = ''.join(df[0].tolist())
print(news_corpus)

# 명사 키워드를 추출
nouns_tagger = Okt()
nouns = nouns_tagger.nouns(news_corpus)
count = Counter(nouns)

# 한글자 키워드를 제거
remove_char_counter = Counter({x : count[x] for x in count if len(x) > 1})
print(remove_char_counter)

# 단어 빈도수가 5 이하인 것들 제거.
df = Counter({key : value for key, value in remove_char_counter.items() if value >= 5})
print(df)
```

빅데이터 - 박성준

텍스트 데이터 마이닝



음성데이터
Wav 파일

AI 학습용
음성데이터 > 텍스트 처리

```
text.append(text_value)
intent_index.append(intent)
label.append(labeling)

except:
    pass

# 데이터프레임 생성
df = pd.DataFrame([x for x in zip(text, intent_index, label)])
df.columns = ['text', 'intent', 'label']

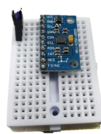
# 데이터프레임 전처리
df = df.rename_axis('index').reset_index()
```

뒷자리 746	영수증번호	4
30분 뒤 도착	소요시간선택	5
가게 도착	가게도착	2
음식점에 왔어	가게도착	2

AI 학습용
텍스트 데이터 전처리

```
[#, 분, 뒤, 도착]      [12, 13, 35, 1]
[가게, 도착]           [8, 1]
```

센서 데이터 마이닝



수집된
센서 데이터
EDA,
이상치 정의

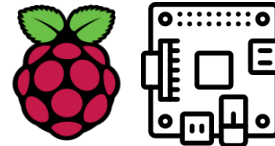
실시간 정규화 이상치 감지

```
# sample1은 직전값 / sample2는 현재값
X_pre_time = (X_sample1 - X_mean) / X_std
X_real_time = (X_sample2 - X_mean) / X_std

detect_X_outlier = abs(X_real_time - X_pre_time)

if detect_X_outlier > X_outlier:
    print("alert")
    count_x += 1
else:
    pass

# Y_sample1은 accel_y 직전값 / Y_sample2는 accel_y 현재값
Y_pre_time = (Y_sample1 - Y_mean) / Y_std
```



배달원 운행 DB 설계

1	2	3	4	5	6
2	3	user3	4	25	100
3	4	user4	9	27	95
4	5	user5	5	58	85
5	6	user6	6	32	85
6	7	user7	0	22	110
7	8	user8	1	46	100
8	9	user9	3	58	95
9	10	user10	7	34	85

배달 데이터 마이닝



이상치
데이터 분석

동 단위 위험감지 파악

```
time_of['time'] = time_of['assignDate'].dt.hour

## 현재 시간 불러오기
i = datetime.datetime.now().hour

# 시간별로 묶기
time_filter = time_of.groupby('time')
time_filter = time_filter.get_group(i)

# 동별로 묶기
dong_time = time_filter.groupby('si_gu_dong')
```

위험감지 지도 시각화

```
# geojson과 동별 인구수를 합친 엑셀 파일 불러오기
final_data = pd.read_excel("./data/geojson/배달데이터_last_final.xlsx")
final_data.columns = final_data.columns.map(str)

# display(final_data.head())

# 동별 구분 경계선
map = folium.Map(location=[37.490, 127.055], zoom_start=12,
```

특정 지역
실시간 위험지도 시각화



빅데이터 - 박성준

보험 데이터 마이닝



수집된
센서 데이터
정규화,
군집분석

군집분석 사전 준비

군집분석 사전 준비

```
# 군집분석 전 yellow bricks 테스트-

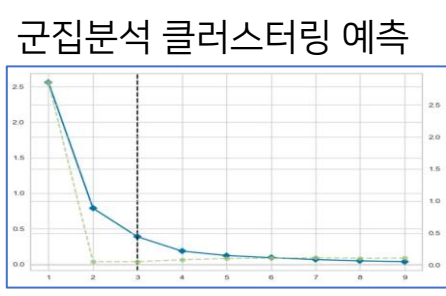
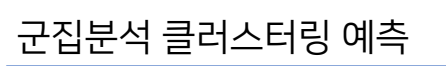
# 데이터 확보
data1 = data[['X변화량', 'Z변화량']]
data2 = data[['X'변화량', 'Y'변화량', 'Z변화량']]

from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler
import pandas as pd

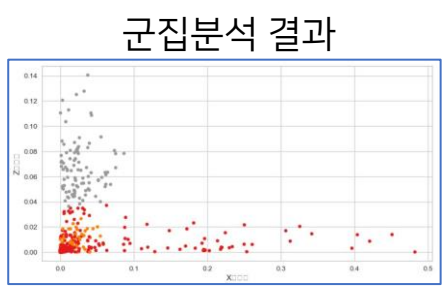
data1 = data1.to_numpy()

from yellowbrick.cluster import KElbowVisualizer

model = KMeans()
visualizer = KElbowVisualizer(model, k=(1,10))
visualizer.fit(data1.reshape(-1,1))
```



군집분석 결과



배달 데이터 마이닝



회귀분석 데이터 확인

```
import pandas as pd

# 데이터 불러오기
df_delivery = pd.read_excel('./data/delivery/서울시일반통계자료.xls')
df_delivery.drop('Unnamed: 0', axis=1, inplace=True)
display(df_delivery.head())

df_classify = pd.read_excel('./data/delivery/서울특별시일반통계자료_201907.xls')
df_classify.drop('Unnamed: 0', axis=1, inplace=True)
display(df_classify.head())
```

변수 시각화



회귀분석 결과

회귀분석 결과

```

from sklearn import linear_model

from sklearn.model_selection import train_test_split

from sklearn.metrics import mean_squared_error

with import opt

# 1. 데이터 불러오기 (데이터셋)
x = np.array([0.01, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 1.0])
x = x.reshape((-1,))

# 2. 데이터 분할 (훈련, 테스트)
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.3, random_state=42)

# 3. 회귀 모델 생성 (선형 회귀)
lr = linear_model.LinearRegression()

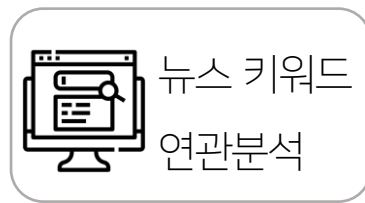
model = lr.fit(x_train, y_train)

# 4. 모델 성능 평가 (R-squared, MSE)
r_squared = model.score(x_train, y_train)
mse = mean_squared_error(y_test, model.predict(x_test))

# 5. 결과 출력
print("R-squared: %.2f" % r_squared)
print("MSE: %.2f" % mse)

```

키워드 연관분석



뉴스기사 크롤링

[illegible]

데이터 전처리

[illegible]

연관분석 결과

연관분석 결과

	support	itemsets
0	0.872727	(배달)
1	0.554545	(안전)

AI- 이동규

텍스트 데이터 전처리

정규표현식으로 숫자치환

Tokenizing

Tokenizer, vocab 사전 조정

빈도순 word_to_index

정수인코딩

Zero 패딩

Simple RNN

	precision	recall	f1-score	support
0	1.0000	0.9630	0.9811	27
1	1.0000	1.0000	1.0000	20
2	0.9737	1.0000	0.9867	37
3	1.0000	0.8824	0.9375	17
4	1.0000	1.0000	1.0000	34
5	1.0000	0.9991	0.9995	11
배달완료	0.7692	1.0000	0.8696	10
accuracy			0.9744	156
macro avg	0.9633	0.9649	0.9610	156
weighted avg	0.9790	0.9744	0.9750	156

```
loss: 0.0816
acc: 0.9834
val_loss: 0.0710
val_acc: 0.9872
```

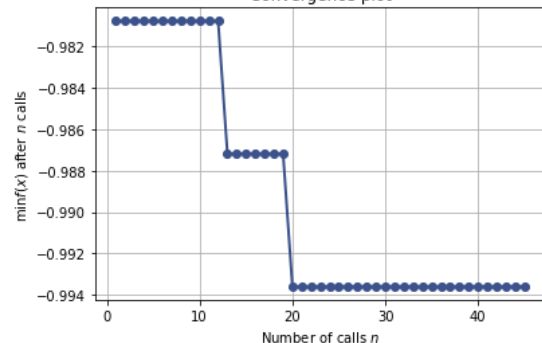
LSTM

	precision	recall	f1-score	support
0	1.0000	0.9630	0.9811	27
1	1.0000	1.0000	1.0000	20
2	1.0000	1.0000	1.0000	37
3	1.0000	0.8824	0.9375	17
4	0.9189	1.0000	0.9577	34
5	1.0000	1.0000	1.0000	11
	1.0000	1.0000	1.0000	10
accuracy			0.9808	156
macro avg	0.9884	0.9779	0.9823	156
weighted avg	0.9823	0.9808	0.9807	156

```
loss: 0.0861
acc: 0.9899
val_loss: 0.0963
val_acc: 0.9936
```

Bayesian optimization

<matplotlib.axes._subplots.AxesSubplot at 0x7fa23b2cf048>
Convergence plot



learning rate: 1.0e-04
num_dense_layers: 2
num_of_dense_nodes: 33
Optimizer: Adam

loss: 0.0861
acc: 0.9909
val_loss: 0.0863
val_acc: 0.9936

Predict 속도 이슈:

실행속도 포함 Bayesian optimization 수행

텍스트 전처리 코드 수정

맞춤법 체크용
Spell_checker 삭제
전처리 시간 50% 개선

최종 모델 선정

learning rate: 4.6e-04
num_dense_layers: 2
num_of_dense_nodes: 512
Optimizer: Adam

loss: 0.0565
acc: 0.9790
val_loss: 0.0270
val_acc: 0.9936
Pred_time: 0.39

IOT - 이주호



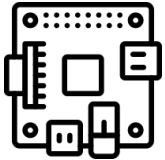
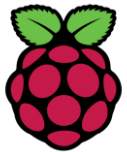
사용자의 wake up word
감지 후 Wav파일 녹음



- 음성을 통한 Application 버튼 처리
- 안전주행 Message 출력



주행 중
별도의 터치 없이
배달대행업무 가능



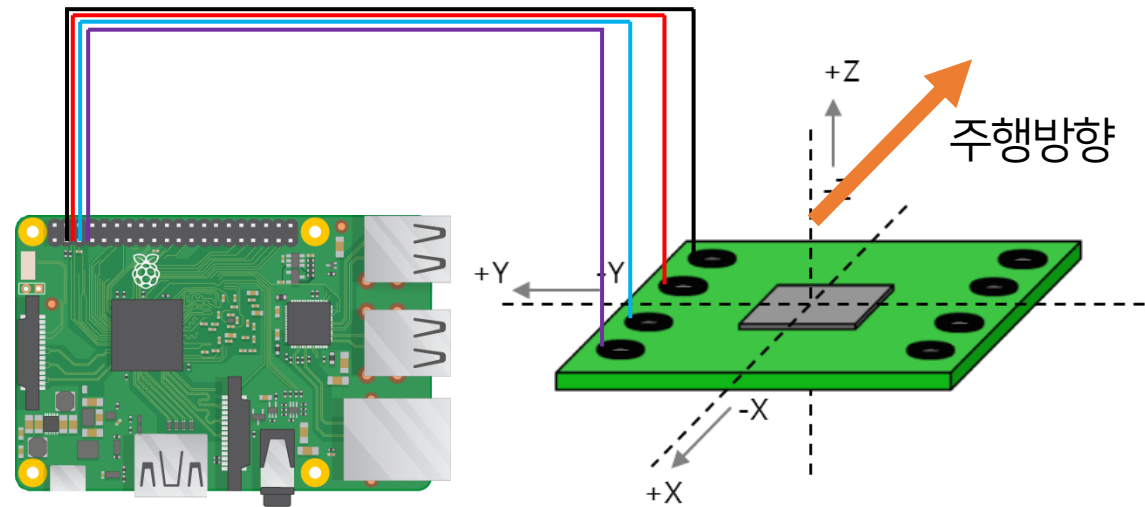
자이로 센서
이상치 감지 & Count

Count값 전송

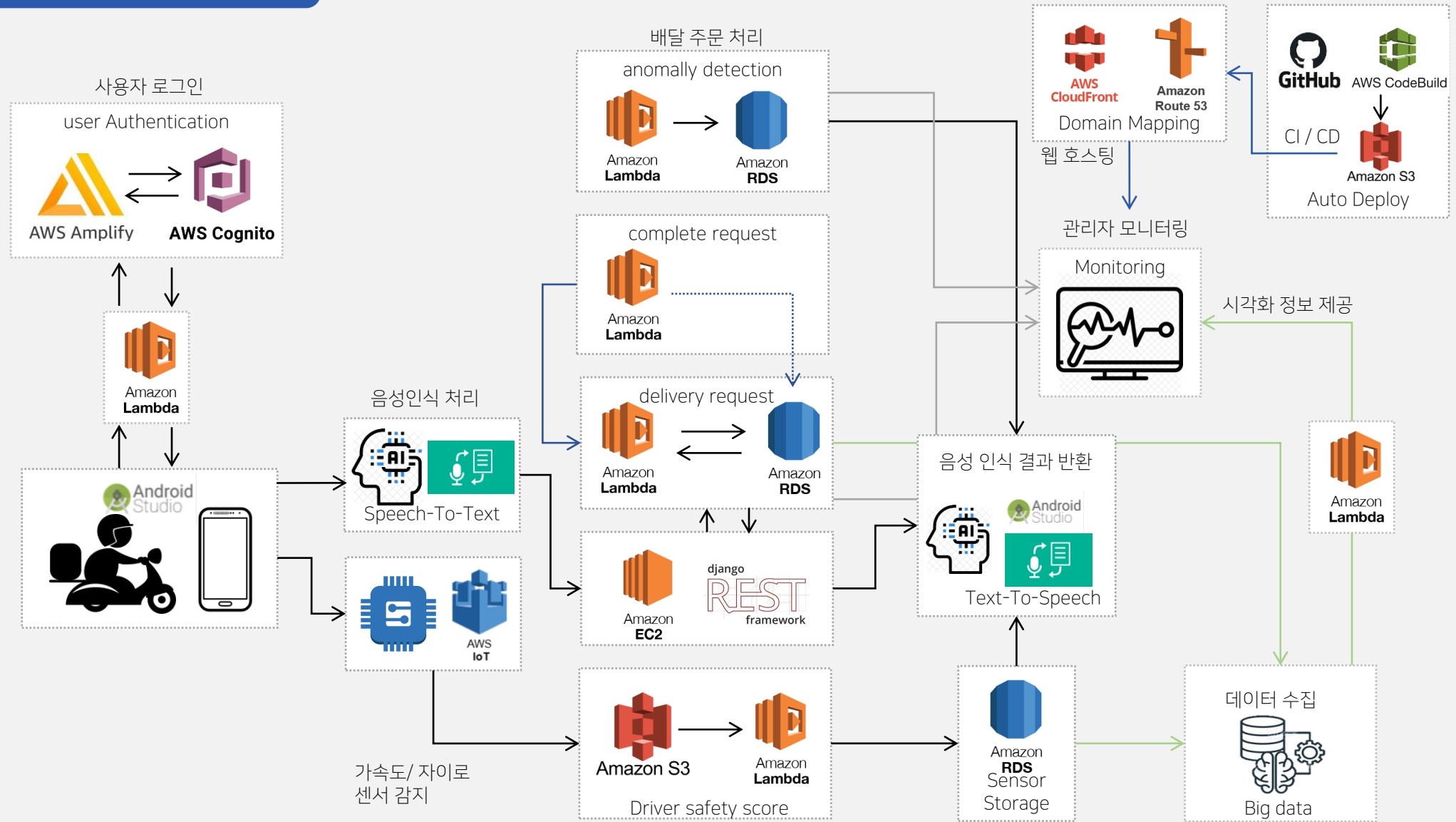
IOT- 이주호



배달 가방에 내장



클라우드 - 이재환

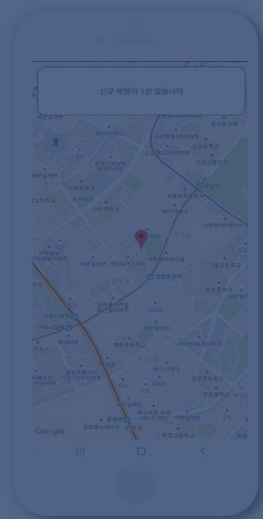


클라우드

작성 API 리스트

idx	Method	API명칭	EndPoint	Request	Response_Success	Define
1	POST	GetLoginUser	https://64mg85mq2c.execute-api.us-east-1.amazonaws.com/hella/login	userid: String	{ "statusCode": "200", "body": "login Success" } { "statusCode": "404", "body": "error" }	Hellacles 사용자 로그인
2	POST	GetLoutoutUser	https://64mg85mq2c.execute-api.us-east-1.amazonaws.com/hella/logout	userid String	{ "statusCode": "200", "body": "logout Success" } { "statusCode": "404", "body": "error" }	Hellacles 사용자 로그아웃
3	GET	GetDeliveryStatus	https://64mg85mq2c.execute-api.us-east-1.amazonaws.com/hella/status	-	{ "statusCode": "200", "status": "status_code" } { "statusCode": "404", "body": "error" }	배달 현황 조회
4	GET	OrderPredict	https://64mg85mq2c.execute-api.us-east-1.amazonaws.com/hella/order	-	{ "statusCode": 200, "predict_order": "predict_order", "pay_avg": "pay_avg", "time_expect": "time_expect" } { "statusCode": "404", "body": "error" }	시간대별 주문 예측 조회
5	GET	getRiderRank	https://64mg85mq2c.execute-api.us-east-1.amazonaws.com/hella/rank	-	{ "statusCode": "200", "body": "rider_rank_list" } { "statusCode": "404", "body": "error" }	배달원 평가 리스트 조회
6	POST	SpeechToText	http://ec2-15-165-62-176.ap-northeast-2.compute.amazonaws.com/SpeechToText	user: String, shopId: Int	{ "statusCode": 200, "body": "intent", "intent_text": "intent_text", "shopName": "shop_name", "destination": "destination" } { "statusCode": "404", "body": "error" }	STT 분석 결과 반환
7	POST	getShopInfo	https://64mg85mq2c.execute-api.us-east-1.amazonaws.com/hella/shopinfo	latitude: Float, longitude: Float	{ "statusCode": "200", "body": "ShopInfoList" } { "statusCode": "404", "body": "error" }	신규 배달 조회

음성인식 챗봇



"음식 픽업 완료"

"네 ** 주문 픽업완료
처리했습니다."

"안전운행 하세요."

For next step

보완점

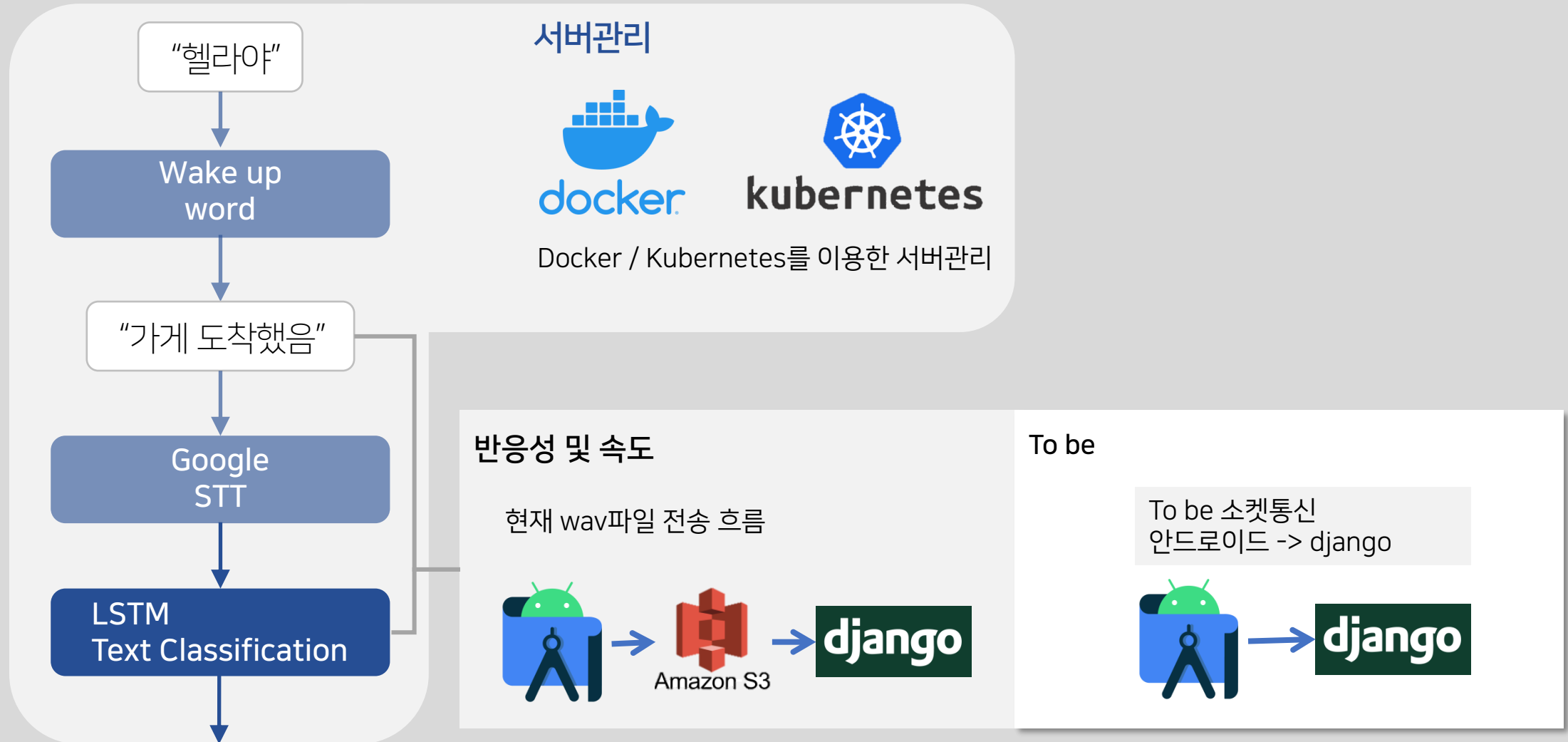
이상치 감지

가속도
자이로 센서



DB: 배달원
운행데이터

보완해야 할 점



보완해야 할 점



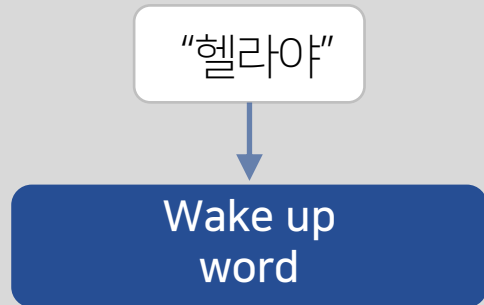
Speech to text 노이즈 이슈

- 야외에서 동작하므로, 소음 제거 이슈
 - 기본 구글 STT가 소음 필터링 기능이 있음. (준수한 성능)
 - WaveNet 등의 소음제거 **pre_trained GAN 모델**이 있으나 모델을 실행하는데 속도 문제
 - 이 부분은 좀 더 스터디 해서 속도와 정확도 측면을 고려해서 추후 개발 할 필요가 있음.
- 성능과 속도가 충분하다면 Wake up word에 적용해서 개발 할 수 있을 것으로 판단됨

긴 문장처리, entity 처리

- 현재는 기존 서비스의 UI를 음성으로 대체한 수준.
- 음성인식은 Deep한 UX를 실현(터치에 비해 과정들을 건너뛰면서)할 수 있다.
 - 이를 활용해 여러 단계의 과정을 한문장으로 입력받고, 이를 한번에 처리할 수 있다면 더 쾌적한 UX구현 가능.
- 예) "헬라야", "**가게 도착**했고 **주문번호 4501**이야"

보완해야 할 점



Wake up word 인식율

- 외부에서 시동어로 음성 인식을 시작하는 경우 음성 인식을 위해 마이크가 항상 ON 상태여야 하는데, 이때 외부 노이즈 제거 등이 제대로 되지 않으면 인식률이 매우 낮습니다.
 - 헬멧 혹은 핸들 쪽에 버튼으로 호출하는 방향으로 변경하여 개발



보험용 센서 데이터 보완

- 현재는 보험용 데이터로 자이로센서 위주의 이상치 감지 수준
 - 추후 센서데이터를 학습해 **사용자의 급커브, 급유턴, 끼어들기 운행 등을 분류하고 count하는 AI 모델**을 만들면 더 정확한 운전데이터 수집 가능할 것으로 생각됨.
 - 가속도 / 자이로 센서 뿐만 아니라 데이터 수집을 위해 다른 센서를 추가하는 것이 좋아보임.

보완해야 할 점



보험용 센서 데이터 보완

- 빅데이터 파트에서의 세분화된 분류 체계 구성
 - AI학습 데이터를 위해 **사용자의 급커브, 급유턴, 끼어들기 운행 등을 분류하고 count할 수 있는 세분화된 분류체계**를 만들면 더 정확한 운전데이터 수집 가능할 것으로 생각됨.
 - 비지도 학습을 통한 구분이 아닌 지도학습을 통해 구분하는 것이 필요.
 - 위험 / 비위험 구분이 아닌 **위험의 정도를 세분화**하는 것이 필요.

배달 데이터 보완

- 예상주문건수 예측을 위한 기타 배달 데이터 수집 필요
 - 주문시간, 연령대 등 이외의 추가적인 데이터를 수집하여 주문건수에 영향을 주는 요인을 알아내기 위한 **회귀분석용 데이터를 수집할 필요**

Drive

운행

Cost

비용

**Easy & Safe
Drive**

운행을 편리하고
안전하게

**Reasonable
Cost**

안전한만큼
비용을 ↓

감사합니다.



Team 헬라클레스

이름	분야	이메일
박성준	데이터 분석 및 엔지니어링	rjs4770@naver.com
이동규	AI 서비스 기획	dklee801@gmail.com
이재환	클라우드 플랫폼 개발	bootre@naver.com
이주호	DBA	justin000@naver.com
홍연하	빅데이터 분석 및 개발	qkdxo102@naver.com