

● 지하철 혼잡도 예측 및 사용자 분산 서비스 기획안

2조- 트랜짓 인사이트 (Transit Insights)

팀장- 이승준

팀원- 김은채 / 이나운 / 이슬 / 최소현 / 허준영

주제 선정 배경

2022년 10월 29일 대한민국의 수도 서울에서 159명 사망, 196명 부상(이태원 참사)

군중밀집 임계점 도달. 출/퇴근길은?

주제 선정 배경

11월 19일

"압사사고 날 것 같아요" 최근길 신고에 강남역 일부 출구 한때 통제

2023.07.21 19:56

서울 강남경찰서 등에 따르면 이날 오후 '강남역에 사람이 너무 많아 **압사 사고**가 날 것 같다.'는 내용의 신고가 접수됐다.

[출처 링크](#)

"지하철 내리면 숨 고르기부터" **일상이 된 과밀**

기사승인 2022-11-01 14:50:29

🔍 📄 🗨️ 📧

2300만 인구가 모여 사는 수도권. **과밀은 일상**이 됐다. 이태원 참사 이후 **다수의 인원이 좁은 공간에 밀집하는 상황**을 개선해야 한다는 지적이 나온다.

[출처 링크](#)

뉴스 > 정책뉴스

지하철 혼잡도 심각하면 무정차 통과...
재난안전문자도 발송

혼잡도 170% 이상 시 '심각' 단계...**철도비상사태** 준하는 조치 시행

[출처 링크](#)

뉴스 > 취재기사

[이태원 참사] 군중밀집 임계점은..."㎡당 5명부터 위험"

출고시간 | 2022-10-31 16:02

스틸 교수는 "신체가 서로 접촉하게 되면, 높은 에너지와 밀도로 인해 **인파가 붕괴**할 수 있다"고 지적했다.

[출처 링크](#)

취재 중에도 2명 실신...“혼잡률 측정, 2년간 3번뿐”

입력 2023.06.17 (06:26) | 수정 2023.06.17 (08:00)

대처 부족, 혼잡률 측정 빈도 부족 등 전체적인 **경각심 필요**

[출처 링크](#)

[르포]"서두르면 쾌적" vs "붐벼도 빨리"...골드라인 대체 버스 1주일, 승자는?

머니투데이 | 김지은 기자

2023.06.28 15:37

🔍 📄 🗨️ 📧

대체 교통이 존재하나, 서비스 안내가 부족해 이용률 저조

[출처 링크](#)

주제 선정 배경

홈 > 사회

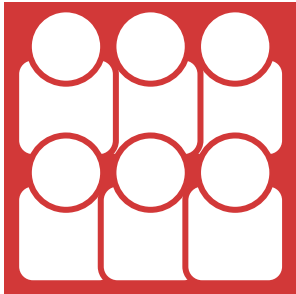
불편 아닌 '안전 문제'된 일상 속 혼잡...1㎡당 6.6명 지하철 어쩌나

[출처 링크](#)

입력: 2022-11-08 17:27 | 수정: 2022-11-08 17:27



1㎡당 6.6명



퇴근시간구로->구일

252%

혼잡도

혼잡도 별 단계



주제 선정 배경

#지하철 #혼잡도 #출퇴근 #교통

호선별 출/퇴근길 가장 혼잡한 구간 [출처링크](#)



장예진 기자 / 201600228

트위터 @yonhap_graphics, 페이스북 tuneey.kr/LeYN1

기존 서비스의 아쉬움

국내에 이런 서비스는 없었을까?

기존 서비스의 아쉬움

지하철 혼잡도를 제공하는 기존 서비스들



티맵 대중교통



지하철 혼잡도 앱



2호선 일부열차

기존 서비스의 아쉬움



서비스 배포 화면

1. 일부 열차의 정보만 서비스 가능
2. 23년 08월 현재 서비스 중단



티맵 대중교통

SK텔레콤에서 운영하는 국내 대중교통 관련 서비스 제공 어플리케이션.

기존 서비스의 아쉬움

1.8



리뷰 11개



Google 사용자



★ ★ ★ ★ ★ 2019년 11월 29일

지하철 칸 수 를 모르겠어요 평균적으로 1-10 역칸 있는데 1 2 3 이렇게 숫자만 있으니 도대체 뭔지 모르겠어요

사용자 1명이 이 리뷰가 유용하다고 평가함

이 리뷰가 유용했나요?

☒ 예

☐ 아니요



Google 사용자



★ ★ ★ ★ ★ 2019년 11월 20일

이름만 혼잡도지 관련정보를 제공하지 않습니다.

사용자 1명이 이 리뷰가 유용하다고 평가함

이 리뷰가 유용했나요?

☒ 예

☐ 아니요



지하철 혼잡도 앱

민간인이 제공하는 지하철 혼잡도 추적
어플리케이션

1. 평균 1.8점, 부정적인 평가가 대다수 - 혼잡 **정보 제공 미비**
2. **서비스 중단** - 마지막 업데이트일 : 2019년 08월 12일

기존 서비스의 아쉬움



1. **최신 열차에서만 확인**이 가능
2. 화면 표시 중 착석 가능 여부를 표현하지 않음



2호선 일부 열차

최신 열차는 각 량에 무게 측정 센서가 붙어 탑승 인원 카운팅이 가능함

기존 서비스의 아쉬움

한 량의 정보에 집중,
역 내 **전체적인 혼잡도 제공 미비**

서비스 제공 시점의 정보만 제공
미래 예측 불가

기대 효과

위 문제들을 해결할 수 있는 서비스가 존재한다면?

기대 효과



안전성 향상

대중교통 혼잡 문제 감소로
이어지는 **안전사고 예방**



편의성 향상

출/퇴근길 불편 해소 및
생활품질 향상



정책 대책

혼잡 예측 및 대응
시스템으로
**신속하고 효율적인
대책 마련**
기반 제공



편익산출

승하차 인원 분산은
직접적인 편익 산출로
이어짐

수행 방향

- 팀 셋업
- 툴 셋업
- 수행 방향
 - 데이터 수집
 - 전처리 / EDA
 - 분석 / 모델개발
 - 예측 / 평가
 - 서비스 배포
- 일정

팀 셋업

역할	이름(가나다순)	세부 역할
팀장	이승준	프로젝트 총괄, 전반작업
팀원 1	김은채	웹개발
팀원 2	이나윤	데이터 수집과 전처리
팀원 3	이슬	데이터 수집과 전처리
팀원 4	최소현	데이터 수집과 전처리
팀원 5	허준영	머신러닝

기초 데이터 수집 - 팀원 전원

주제 선정을 위한 자료와 대략적인 데이터 수집

팀장
—
이승준

세부 데이터 수집과 전처리
이나윤, 이슬, 최소현

머신러닝 : 허준영

웹 개발: 김은채

툴 셋업 - 예상안

ML- 예측

Scikit-learn



XGBoost



LightGBM



웹 - 대시보드

Django



Streamlit



matplotlib



Plotly



웹 - 서비스

Django



Plotly



수행방향

1

데이터 수집

실시간 지하철 혼잡 예측을
위한 데이터 수집.

- . 승차 인원
- . 기상 데이터
- . 휴일/ 공휴일 날짜
- . 역사별 혼잡도

2

전처리 / EDA

각각의 독립 변수(시간,
날씨, 공휴일 등)와
종속변수(승하차인원)
의 상관관계 분석 및
시각화

3

분석 / 모델개발

회귀분석과
분석 모델 개발

4

예측 / 평가

평가 지표와
라이브러리를 활용,
모델의 적합성 판단

5

서비스 배포

시민들이 어디서든
지하철 혼잡도를
확인할 수 있도록
서비스 개발 및 배포

수집 데이터 개요

분류	생성 변수	데이터 명	출처	비고
지역	지하철 위치	자치구별지하철역정보_20230531.CSV	https://www.data.go.kr/data/15081868/fileData.do	공공데이터 OpenAPI
지역	동 IDName	서울시 읍면동 마스터 정보	https://data.seoul.go.kr/dataList/OA-21234/S/1/datasetView.do	서울 열린데이터 OpenAPI
승객	시간별 혼잡도	서울시 주요 113 장소 실시간 데이터	https://data.seoul.go.kr/dataList/OA-21285/A/1/datasetView.do	서울 열린데이터 OpenAPI
승객	시간별 승/하차 인구	서울시 지하철 호선별 역별 시간대별 승하차 인원 정보	https://data.seoul.go.kr/dataList/OA-12252/S/1/datasetView.do	서울 열린데이터 OpenAPI
기상	일일 기상 조건	[기상_101]_날짜_요일_기상조건	기상청	일일 기온 및 강수량 기준
시간	지하철 시간 데이터	서울시 지하철 실시간 도착 정보	https://data.seoul.go.kr/dataList/OA-12764/F/1/datasetView.do	서울 열린데이터 OpenAPI
시간	21~23 대한민국 공휴일	공휴일 21년~23년.csv 크롤링_공휴일_수정2.ipynb	https://www.data.go.kr/tcs/dss/selectApiDataDetailView.do?publicDataPk=15012690	공공데이터 openAPI

가설 설정

독립변수

시간대	요일	기상
출.퇴근 새벽 오전	평일 토요일 공휴일	기상 상황



독립변수로 시간대, 요일, 기상조건 설정,
지하철 혼잡도에 영향을 미칠 것으로 예상

승객 수 1량(명)	혼잡도 환산
54	34%
94	59%
160	100%
240	150%
368	230%

데이터 분석 프로세스

- 1 데이터 추가 수집 / 추가 필요한 자료 수집
- 2 전처리 / 결측치 처리, 필요 데이터 추출, 분할
- 3 변수 선정 / EDA를 통한 상관관계 분석, 파생변수 생성 후 의미있는 변수만 선택
- 4 머신러닝 / 변수 특성에 적합한 학습 모델 선정 및 모델 학습(fit)
- 5 시각화 및 검증 / 분석 결과 시각화 및 가설검증을 통한 머신러닝 모델 검증

분석 모델 개발 & 평가 및 예측

1

회귀 모델 후보 선정: 단순선형회귀 or 다중선형회귀: LinearRegression, LGBMRegressor

2

데이터셋 분리 및 학습: 훈련: 80%, 테스트: 20%, 교차검증횟수: N(분석에따라다름)

3

평가 지표: MAE, RMSE, R2 score

4

최종 모델 선정 및 데이터 시각화: 시간-예측혼잡도 그래프(예시)

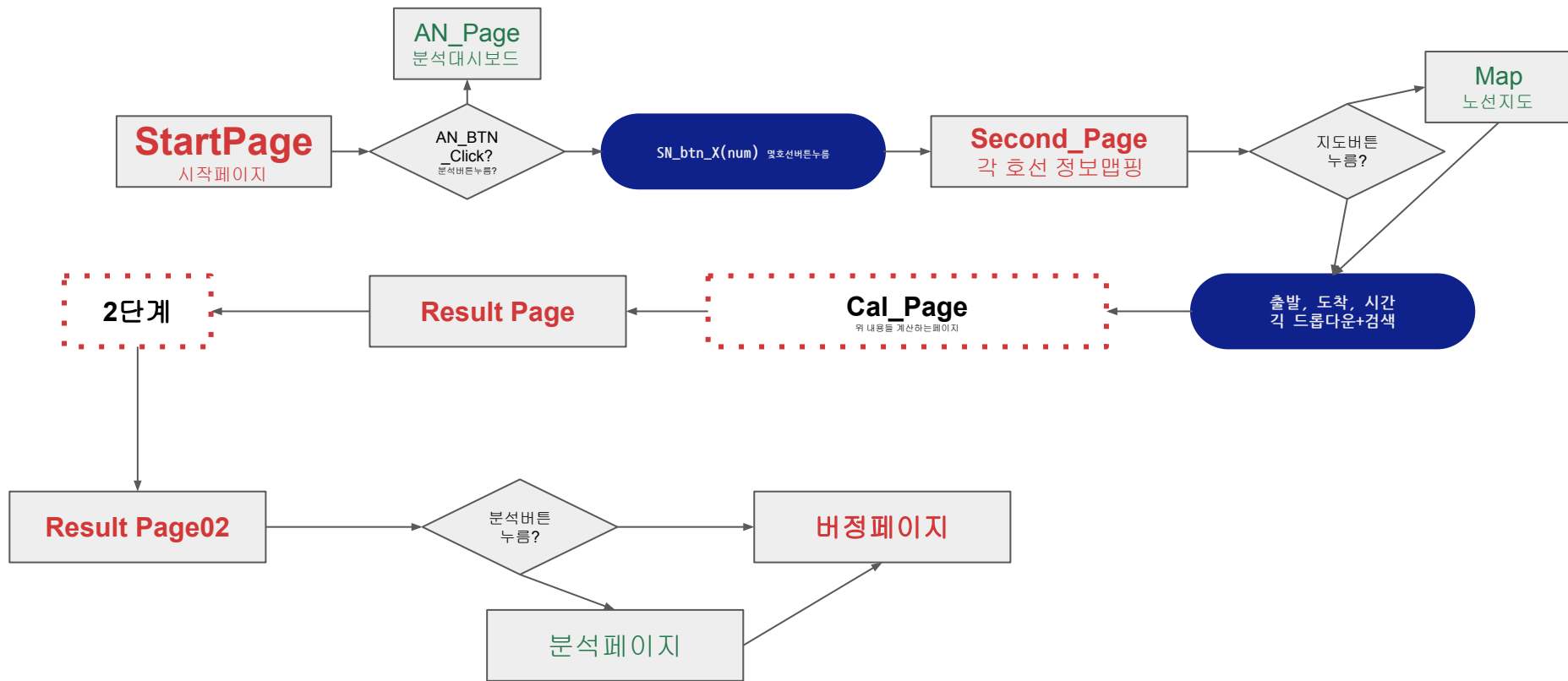
서비스 구현(예상안)

<https://youtu.be/hcHH2xDA7cc>

The screenshot shows a web browser window with a URL bar containing "https://multimedia2019.Team-Z.com". The application interface is divided into several sections:

- Left Sidebar:** A vertical list of buttons labeled "1호선" through "9호선" in various colors. Above this list, it says "기온 28°C".
- Top Center:** A dark blue button with white text that reads "클릭하여 1호선 노선도 확인하기".
- Form Area:** Three input fields are stacked vertically: "서울역", "용산역", and a time range "14:00~15:00" with a dropdown arrow. A blue "검색" (Search) button is located below the time field.
- Result Card:** A large white rounded rectangle containing the following information:
 - Route: "서울역 => 용산역"
 - Time: "14:00~15:00"
 - Text: "예상 혼잡도는"
 - 120% 착석불가** (Large bold text)
 - Subtext: "5단계 중 2단계"
 - Footnote: "다른 방법은 여하신가요?"

서비스 구현(예상안)

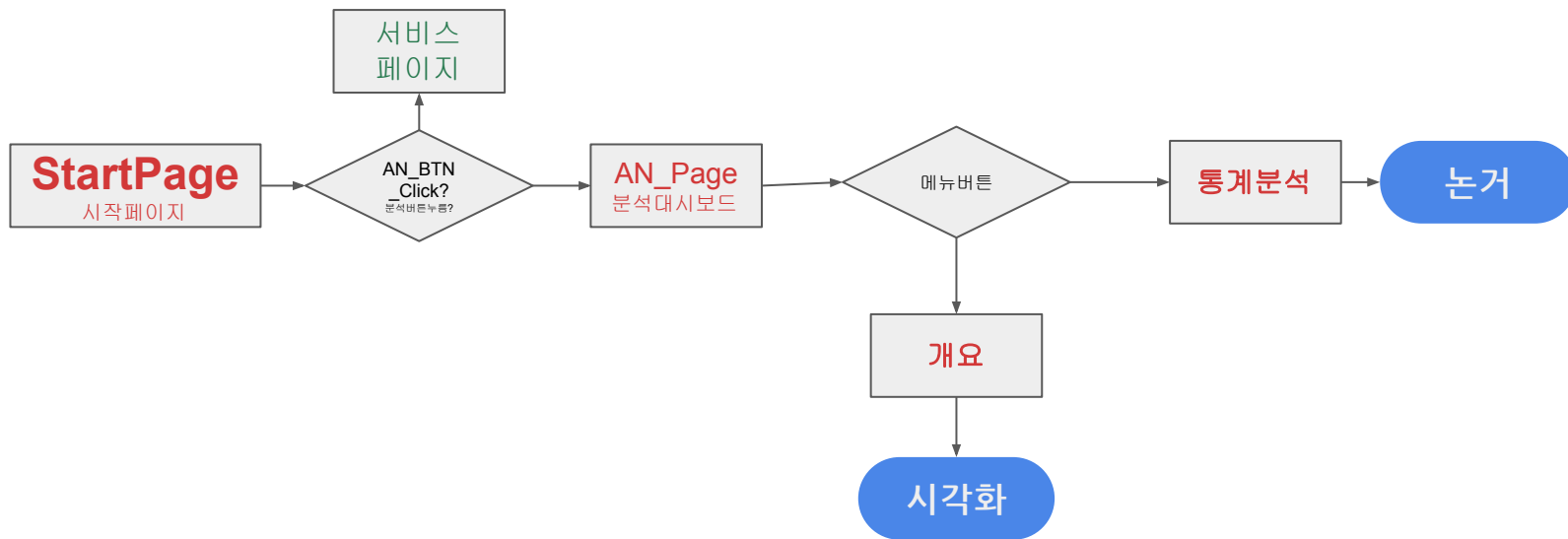


대시보드 배포(예상안)

<https://youtu.be/b9quYNI1uoU>



대시보드 배포(예상안)



[illegible]

전체 일정

일	월	화	수	목	금	토
					09/1	2
					기획	
3	4	5	6	7	8	9
기획				데이터 전처리		
데이터 수집						
10	11	12	13	14	15	16
데이터 전처리				분석 / 모델평가		
EDA / 시각화						
17	18	19	20	21	22	23
분석 / 모델평가						
EDA / 시각화						
24	25	26	27	28	29	30
전체 마무리		발표!				

출처

뉴스(제목링크) —

[해럴드 경제: “압사사고 날 것 같아요” 퇴근길 신고에 강남역 일부 출구 한때 통제\(링크\)](#)

[쿠키뉴스: “지하철 내리면 숨 고르기부터” 일상이 된 과밀](#)

[대한민국 정책 브리핑지하철 혼잡도 심각하면 무정차 통과…재난안전문자도 발송](#)

[서울신문: 불편 아닌 ‘안전 문제’된 일상 속 혼잡…1㎡당 6.6명 지하철 어쩌나](#)

[퍼즐: 호선별 출/퇴근길 가장 혼잡한 구간](#)

논문(제목링크) —

[기상에 따른 대중교통 이용변화의 영향도 분석](#)

Q&A