

공공 빅데이터 융합을 통한 서울 지하철 혼잡도 예측 및 설명가능한 AI 기반 정책 대응

Jaeheung Park^a, Kyungwon Kim^{a,*}

^aSchool of International Trade and Business, Incheon National University, Incheon, 22012, Republic of Korea

Abstract

Background: 전세계에서 대중교통 이용률이 가장 높은 나라는 대한민국이다. 특히 대한민국의 수도인 서울의 지하철은 약 600여개의 역사에서 매일 약 800만명이 이용할 만큼 대표적인 교통 수단이다. 2022년 기준 하루 평균 대중 교통 이용건수는 약 1,025만 건인데, 그 중에서 약 50.3%를 차지할 정도로 서울시민 중 절반 이상이 매일 지하철을 이용한다. 교통시설이나 수단, 그리고 이용대상에 따라 관련 이론이나 방법들이 개발되고 있지만, 설명가능한 인공지능을 사용하여 혼잡도를 정교하게 예측하고 그 근거를 제시하는 연구는 거의 전무한 실정이다.

Objective: 본 연구에서는 시민들이 객차 내외부에서 느낄 수 있는 승강장 내 혼잡도를 예측하는 것을 목적으로 한다. 또한 실시간으로 정확하게 혼잡도를 예측 및 설명함으로써 빅데이터 기반 비즈니스 애널리틱스 방법론을 제시하고 혼잡도 원인에 대한 대응을 효과적으로 제시한다.

Participants and Methods: 대한민국의 수도인 서울의 승강장 내 혼잡도를 예측하기 위해 총 8개의 공공데이터포털의 17개 데이터베이스를 융합하였다. 그리고 대표적인 머신러닝 및 딥러닝 알고리즘으로 혼잡도를 예측하였으며, 설명가능한 인공지능인 SHAP 알고리즘으로 예측된 혼잡도의 근거를 실시간 효과적으로 제시하였다.

Results: 대표적 머신러닝 알고리즘은 Random Forest를 사용하여 99% 이상의 미래 혼잡도 예측 성능을 달성하였으며, 청소년이나 노인층의 탑승 비율을 낮추거나 경기 선행지수나 실업율 등이 낮아질 경우 지하철의 혼잡도를 낮출 수 있음을 확인하였다. 또한 상행선의 평균운행간격을 높이거나 기준금리, 달러환율, 경기 동행지수나 지하철 요금이 높아질 경우 혼잡도를 역시 낮출 수 있었다.

Conclusions: 설명가능한 인공지능을 사용한 비즈니스 애널리틱스 방법론은 실시간으로 대한민국 수도인 서울의 혼잡도를 예측할 수 있을 뿐만 아니라 선제적으로 원인을 파악하여 혼잡도를 낮추기 위한 대중교통이나 도시개발 정책에 대응하는데 도움을 줄 수 있습니다.

Keywords: Subway Congestion, Explainable Prediction, Machine and Deep Learning, SHAP, Business Decision

*Corresponding author: thekimk.kr@gmail.com
ORCID(s): 0000-0001-6530-8426 (K. Kim)

1. Introduction

대중교통은 도시의 경제적 활력과 사회적 지속가능성을 지탱하는 핵심 인프라로서, 시민들의 일상과 밀접하게 연계되어 있다. 글로벌 시장조사기관 Statista의 최근 분석에 따르면, 전 세계에서 대중교통 이용률이 가장 높은 국가는 대한민국으로 보고되었다 (Fleck 2023). 통근·통학 시 대중교통을 활용한다는 응답률은 41%에 달했으며, 특히 지하철은 전체 대중교통 중 가장 높은 30%의 수단 분담률을 기록하였다. 그중에서도 서울의 지하철은 약 600여 개 역사에서 매일 800만 명 이상이 이용하는 대표적인 교통 수단으로, 2022년 하루 평균 대중교통 이용 건수 1,025만 건 가운데 절반 이상(50.3%)을 차지하였다 (Office 2023)[1].

서울은 주거 및 업무 중심 지역의 밀집과 더불어, 출퇴근 시간대와 사회·정치적 이벤트로 인해 특정 시기의 지하철 혼잡이 구조적으로 심화되는 양상을 보인다. 혼잡으로 인한 사회적 비용은 막대하다. 지하철 내 혼잡으로 인한 비용은 약 2,317억 원, 열차 지연에 따른 비용은 4,960억 원으로 추산된다 (김승준 2016)[2]. 이에 대응하기 위해 서울시는 역사 내 물리적 시설 확충에 주력해왔으나, 수요 분산 효과는 제한적이었다. 향후 신도시 개발 및 새로운 교통수단 도입에도 불구하고, 전 시간대에 걸친 혼잡 해소 가능성은 불확실하다. 더 나아가, 코로나19 팬데믹은 밀폐된 대중교통 공간에서의 전염 위험성을 극명히 보여주며 혼잡 문제가 단순히 교통 효율성의 문제가 아니라 보건·사회적 위기 관리와도 직결됨을 시사하였다.

따라서 다수의 시민이 매일 이용하는 지하철의 혼잡도를 사전에 정밀하게 예측하는 것은 단순한 운영 효율성 차원을 넘어, 시민들의 안전·편의성 제고와 도시 교통정책 수립의 핵심 과제로 자리매김한다 (이상준 2020)[3]. 혼잡도를 높은 정확도로 예측할 수 있다면, 이는 곧 승객 흐름과 이용 패턴의 실시간 파악을 가능케 하며, 운행 빈도 조정·열차 편성·인력 배치 등 운영 전략의 최적화에 기여할 수 있다. 또한 도시계획 및 지속가능한 교통 인프라 정책의 근거 자료로 활용될 수 있으며, 응급 상황에서의 대응 역량을 강화함으로써 시민 안전을 보장하는 기반이 된다 (이선하 et al. 2015)[4].

교통상태의 혼잡을 완화하는 연구는 지속적으로 해결하려는 난제이다. 교통 시설이나 수단, 그리고 이용대상에 따라서 관련된 이론이나 방법론들이 다양하게 개발되고 있다. 이호 and 최진경 (2015)은 승강장 혼잡도를 승강장 면적 대비

승차대기인원으로 정의하였다. 이 때 승차대기인원은 열차 탑승을 위해 대기하는 인원으로 열차 출발정보와 승강장 누적진입인원을 활용하였다 [5]. 이상준 et al. (2021)은 승강장 내 승객수와 승객점유면적의 비율을 사용하여 혼잡도를 정의하였다. 승강장 내 승객이 실제 이용하는 대기 면적 개념을 제안하였으며, 승차승객 대기면적계수를 추정하였다 [6]. 이상준 (2020)은 1분 단위 통행량 데이터를 기반으로 실용 대기 면적을 산정하고, 이를 토대로 승강장 동적 혼잡도를 추정하는 방법론을 제시하였다 [7]. 이러한 연구들은 혼잡도를 정량화하는 기초적 토대를 마련했으나, 여전히 복잡한 요인들이 실시간으로 혼합적으로 작용하는 실제 혼잡 현상을 설명하고 예측하기에는 한계가 존재한다.

기존 연구들을 종합하면, 지하철 혼잡도는 크게 세 가지 차원으로 구분할 수 있다. 첫째는 승강장에서 승하차를 기다리는 과정에서 발생하는 혼잡도이며, 둘째는 열차에 탑승한 이후 객차 내부에서 승객이 체감하는 혼잡도이다. 예컨대 공공데이터포털의 “서울교통공사_지하철혼잡도정보” 데이터베이스에서는 혼잡도를 정원 대비 승차 인원으로 정의하고, 좌석 수와 승차 인원이 일치할 경우 혼잡도를 34%로 산정한다. 마지막으로, 승강장 대기 혼잡도와 객차 내 혼잡도를 모두 포괄하여 이용자가 전체 여정에서 경험하는 총체적 혼잡도로 이해할 수 있다. 특정 연구가 객차 내부만을 고려한다면 승강장에 대기하는 승객들의 피로와 불편을 반영하지 못할 수 있으며, 반대로 객차 외부만을 고려한다면 사회적 요인으로 객차는 비교적 여유로우나 승강장 자체가 혼잡한 상황을 설명하기 어렵다. 따라서 보다 포괄적 관점에서 승강장과 객차를 아우르는 실질적 혼잡 경험을 정의할 필요가 있다.

본 연구는 이러한 맥락에서 시민이 객차 내부와 외부에서 동시에 체감할 수 있는 모든 혼잡 상황을 반영하기 위해, 면적 대비 승·하차 인원을 핵심 지표로 설정하여 예측을 수행한다. 최근 들어 혼잡도 연구는 전통적인 통계 기반 접근을 넘어 머신러닝과 딥러닝 기법을 활용하는 방향으로 확장되고 있다. 예를 들어, 김재익 (2013)은 지하철 승·하차량과 주변 토지이용 특성과의 관계를 분석하여, 주거지·학교·상업시설의 비율이 높을수록 하차 수요가 증가함을 확인하였다. 김진수 (2016)는 SNS 데이터를 포함한 다중회귀모형을 적용했으나 예측 성능은 제한적이었다. 성현곤 (2017)은 6년간의 대중교통 이용 데이터를 분석하여, 인구 감소에도 불구하고 철도망 확충과 서비스 개선이 이용 증가로 이어진다는 점을 규명하였다. 또한 이정훈 and 정현영 (2018)은

분위 회귀모형을 활용해 기온·습도·풍속 등 기상 요인이 대중교통 수요와 밀접히 연관됨을 밝혔다 [10].

머신러닝과 딥러닝 기법을 적용한 연구는 성능 측면에서 더욱 고무적인 성과를 보이고 있다. 예컨대 이경재 et al. (2020)은 대표적인 딥러닝 알고리즘인 다층퍼셉트론을, 권효승 et al. (2020)은 순환신경망의 일종인 GRU 알고리즘을 적용하여 지하철 구간 혼잡도를 30분 단위로 예측하였고, 낮은 손실값(0.0431)을 기록하며 예측력 향상을 입증하였다. 이는 지하철 혼잡도를 GRU 기법으로 예측 모델을 만들었다는 점에서 의의가 있다. 장진영, 김채원, and 박민서 (2023)는 로지스틱 회귀와 더불어 의사결정나무 및 랜덤포레스트(Random Forest)를 활용하여 80% 이상의 정확도를 달성했으며, 버스정류장 수나 역세권 건물 수 등 새로운 예측 지표를 도입한 점에서 의의가 있다 [14]. 최근 김규진 et al. (2024)은 공공데이터포털, 기상자료개방포털, 열린데이터광장 등 다양한 공공 빅데이터를 융합하고 다층 퍼셉트론을 적용하여 시간대별·기상 조건별 혼잡도를 정밀하게 예측하였다. 시간대별 특성과 기상 조건을 고려하여 지하철역 혼잡도를 예측한 점에서 의의가 있다 [17]. 이와 더불어 버스나 도로 교통 상황까지 확장하여 교통 도메인 특성에 맞춘 딥러닝 알고리즘들이 지속적으로 개발되어 높은 성능을 달성한 연구들도 많아지고 있다 [20-21]. 그 외에 관련 선행연구는 <표 1>의 연구목적 별로 상세하게 확인할 수 있다. 전반적으로 머신러닝과 딥러닝 알고리즘을 활용한 혼잡도 예측 연구가 성능이 증가되고 있음을 알 수 있다.

<표 1> 지하철 혼잡도 예측과 교통분야의 인공지능과 XAI를 활용한 선행연구 분석 방향과 성능 요약

연구목적	선행연구	데이터 소스	데이터 기간	종속변수	분석내용 및 방법론	성능
정량적 지하철 혼잡도 예측	김근원 외(2015) [8]	서울시열린데이터광장, 서울시통계 (3 개 DB 활용)	2013 년	시간대별 객차 내 정원대비 인원	1~4 호선 대상 상관관계 네트워크 1 개	시각화 분석
	Kim(2016) [9]	국가통계포털, SNS 크롤링 (3 개 DB 활용)	2013 ~ 2014 년	시간대별 객차 내 정원대비 인원	1 호선 대상 다중회귀분석 1 개	다중회귀분석 기준 75~84%
	Lee 외(2019) [10]	기상청자료, 교통카드자료 (DB 수 미기재)	2011 ~ 2015 년	연도별 승객통행량	대중교통 대상 분위회귀분석 1 개	변수별 통계수치 분석
	Shin 외(2019) [11]	교통카드자료 (1 개 DB 활용)	#####	시간대별 면적 대비 승강장통행량	1~9 호선 대상 규칙기반 알고리즘 1 개	변수별 통계수치 분석
	Liu 외(2019)	Nanjing Metro System	2016년 3 월 18 일 ~ 4 월 30 일,	승객 경로	단정 1 개역 대상 ARIMA, 딥러닝 등	딥러닝 기준 RMSE 8.19

	[12]	(1 개 DB 활용)	2016년 8월 1일 ~ 11월 9일		10 개	
권효승 외(2020) [13]	서울시 빅데이터 캠퍼스 (3 개 DB 활용)	2019년 10월 ~ 2020년 7월	시간대 및 최단경로별 승하차인원	1~9 호선 대상 GRU 1 개	GRU 기준 Loss 0.0431	
Lee 외(2020) [7]	교통카드자료 (1 개 DB 활용)	#####	시간대별 면적 대비 승강장통행량	모든노선 대상 규칙기반 알고리즘 1 개	변수별 통계수치 분석	
Lee 외(2021) [6]	교통카드자료 (1 개 DB 활용)	#####	승객점유면적 대비 승차대기인원	통행량 상위 30 개역 대상 규칙기반 알고리즘 1 개	변수별 통계수치 분석	
Jang 외(2023) [14]	공공데이터포털 (4 개 DB 활용)	2019, 2022년	시간대별 객차 내 정원대비 인원	2 호선 대상 Logistic Regression, Decision Tree, Random Forest 3 개	Random Forest 기준 83~99%	
Park 외(2023) [15]	공공데이터포털 (1 개 DB 활용)	미기재	시간대별 객차 내 정원대비 인원	3,5,6 호선 대상 그래프 기반 딥러닝 1 개	그래프 딥러닝 기준 34~97%	
Hu(2023) [16]	Subway Card Swipe Data (1 개 DB 활용)	2019년 1월 1일 ~ 20일	승객 유입과 유출	항저우 3 개노선 대상 LSTM 1 개	LSTM 기준 R^2 97.1%	
Kim 외(2024) [17]	공공데이터포털, 서울시열린데이터광장, 기상자료개방포털 (DB 수미기재)	2021 ~ 2023년	시간대별 객차 승하차인원	1~8 호선 대상 MLP 1 개	MLP 기준 84%	
Tian 외(2024) [18]	EZ-Link Card Data (1 개 DB 활용)	2016년 1월 ~ 3 월	승객수요	싱가폴 7 호선 대상 ARIMA, Random Forest, LSTM, 자체개발 4 개	자체개발 기준 MSE 2433.20	
Luo 외(2025) [19]	자체수집 (DB 수 불필요)	2020년 1월 ~ 2021년 3 월	시간대별 객차 내 정원대비 인원	부산 1 호선 대상 RNN 및 그래프 기반 딥러닝 6 개	Conv-LSTM 기준 89.11%	
교통분야 AI 기반 혼잡도 예측	Li 외(2018) [20]	METR-LA of Los Angeles Country, PEMS-BAY of California (2 개 DB 활용)	2012년 3월 1일 ~ 6월 30일, 2017년 1월 1일 ~ 5월 31일	교통 속도	ARIMA, VAR, SVR, FNN, FC-LSTM, DCNN, DCRNN, DCRNN-SEQ, GCRNN 등 10 개	DCRNN 기준 89.5~97.1%
	Pei 외(2023) [21]	IC Card Swipe Data (1 개 DB 활용)	2021년 3월 ~ 10 월	버스 승객수	SVR, XGBoost, Bi-LSTM, ATT-BiLSTM, WPD-BiLSTM, WPD-ATT-BiLSTM 6 개	WDP-ATT-BiLSTM 기준 91.453~97.717%
교통분야 XAI 기반 설명	Monje 외(2022) [22]	IC Card Swipe Data (1 개 DB 활용)	2015년 1월 ~ 2017년 2 월	버스 승객수	XGBoost, Random Forest, SVM, DNN, LSTM 등 7 개	LSTM 기준 R^2 89%
	Lee(2023) [23]	도시재생종합정보체계, 한국교통연구원, 통계청 (DB 수 미기재)	2019년	인구밀도와 접근성	3000 여개 읍면동 대상 머신러닝 + XAI 7 개	XGBoost 기준 R^2 91.3%
	홍석준 외(2024) [24]	도로와 차량 드론 촬영 데이터	2024년 1 주일	차선변경여부	대전시 유성구 대상 LSTM, Transformer, 자체개발 + XAI 3 개	자체개발 기준 94.73~95.46%
	Lee 외(2024) [25]	SKT 이동통신자료	2024년 1 주일	화물과 비화물 운전자 통행분류	딥러닝 + XAI 1 개	딥러닝 기준 71~83%

Aboulola(2024) [26]	Accident Data from New Zealand Crash Analysis System (2 개 DB 활용)	2016 ~ 2020 년	사고 부상 심각도	MLP, CNN, LSTM, ResNet, EfficientNetB4, InceptionV3, Xception, MobileNet 8 개	MobileNet 기준 F1-score 98.48%
------------------------	---	---------------	-----------	--	---------------------------------

요컨대, 기존 연구들은 통계적 접근부터 최신 딥러닝 기반 모델까지 폭넓게 시도되었으며, 전반적으로 예측 성능이 점차 고도화되는 추세임을 보여준다. 그러나 여전히 승강장과 객차라는 이분법적 구분에 머무르는 경향이 있어, 시민들이 실제로 체감하는 총체적 혼잡 경험을 설명하기에는 한계가 존재한다. 본 연구는 이러한 공백을 메우고자, 공공 빅데이터 융합과 설명가능한 인공지능을 결합하여 지하철 혼잡도를 보다 실질적이고 해석 가능한 방식으로 예측하고자 한다.

머신러닝과 딥러닝은 이미 다양한 교통 예측 분야에서 탁월한 성능을 보여주고 있으나, 그 의사결정 과정이 불투명하다는 점에서 흔히 “블랙박스(Black Box)” 모델로 지칭된다. 즉, 높은 정확도의 예측값은 제공하지만, 그 근거를 사람이 직관적으로 이해하거나 검증하기 어렵다는 태생적 한계가 존재한다. 이러한 한계를 극복하기 위해 최근 학계와 산업체에서는 설명가능한 인공지능(XAI)에 대한 관심이 급격히 고조되고 있다. XAI는 단순히 “무엇을 예측했는가”를 넘어 “왜 그렇게 예측했는가”를 설명할 수 있도록 지원하며 [27], 이는 모델 개발자뿐 아니라 정책 입안자, 서비스 관리자, 최종 이용자에게도 중요한 해석적 단서를 제공한다. 결과적으로, XAI는 인공지능 모델의 신뢰성(trustworthiness)과 투명성(transparency)을 제고하는 핵심 도구로 부상하고 있다 [28].

지하철 혼잡도 분야에서는 아직까지 XAI의 적용이 거의 전무하지만, 다른 교통 연구 영역에서는 인공지능 기반 예측과 설명을 결합한 시도가 점차 확산되고 있다. 예컨대, 버스 승객 수를 예측하거나 도시 내 인구밀도를 분석하는 연구 뿐 아니라 [22-23], 교통 상황 및 사고 발생 요인을 해석하려는 연구가 활발히 이루어지고 있다 [24-26]. 이는 XAI가 단순히 예측 정확도를 높이는 차원을 넘어, 인간과 인공지능 간의 상호작용을 개선하고, 책임 있는 AI 활용(responsible AI use)을 촉진한다는 점에서 학문적·실무적 의의가 크다는 것을 시사한다.

물론, XAI 역시 한계가 존재한다. 분석 후 추가 계산이 요구되며, 일부 방법론은 전체(global) 설명보다는 부분적(local) 설명에 머물고, 설명 결과가 비전문가에게 다소

난해할 수 있다. 또한 모든 데이터와 모델에 보편적으로 적용되기 어렵고, 복잡한 변수 간 상호작용을 해석하는 데 기술적 제약이 따른다. 그럼에도 불구하고, 이러한 한계를 극복하려는 연구는 점차 확산되고 있으며, AI 예측 성능과 해석 가능성은 동시에 확보하는 접근은 교통 연구와 정책 설계에 새로운 가능성을 열고 있다.

대표적인 XAI 기법으로는 반사실적 설명(Counterfactual Explanations), 변수 중요도(Feature Importance), LIME(Local Interpretable Model-agnostic Explanations), 그리고 SHAP(Shapley Additive Explanations) 등이 있다. 이 중에서도 SHAP은 게임이론에 기반하여 각 변수의 기여도를 직관적이고 일관되게 산출할 수 있어 기술적 완성도가 높다. 본 연구는 SHAP을 활용하여, 단순히 고성능의 지하철 혼잡도 예측 모델을 구축하는 데 그치지 않고, 예측 결과의 근거를 명확히 제시하고 변수 기여도를 정량적으로 분석한다. 이를 통해 기존 연구가 제공하지 못했던 설명 가능성과 실무적 해석력을 동시에 확보하고, 나아가 지하철 혼잡도 관리 및 정책 대응을 위한 새로운 학문적·실무적 기여를 제시하며 기존 문헌들을 확장하였다.

본 연구는 기존 선행연구와 달리 다음과 같은 5가지 항목에서 기여하고 있다. 첫째, 2024년 현재 지하철 관련 공공 빅데이터를 모두 융합하여 혼잡도 예측에 활용하였다. 기존 연구들은 적은 수의 공공데이터만을 활용했지만 본 연구는 총 17개의 데이터베이스를 융합하여 공공데이터의 효용을 뒷받침할 수 있었다. 둘째, 단순히 지하철 관련 데이터베이스 뿐만 아니라 국내외 경제 관련 데이터를 반영하여 지하철 혼잡도와의 관련성을 규명할 수 있었다. 셋째, 기존 연구들이 비즈니스적으로 바로 활용되기엔 낮은 예측 성능이었지만, 본 연구에선 MSPE 기준 99% 이상의 예측 정확성을 보이며 대중교통 관계자들이 실시간으로 바로 활용할 수 있는 방법론을 달성하였다. 이러한 접근은 다양한 시간대와 노선, 그리고 여러 기상 상황에서도 혼잡도를 효과적으로 예측함으로써 효율적인 자원 배분과 운영 계획 수립에 도움될 수 있다. 넷째, 높은 예측 성능 뿐만 아니라 설명가능한 인공지능을 활용하여 예측 결과를 보다 쉽게 이해할 수 있도록 제시하였다. 설명가능한 인공지능을 통해 예측 결정 과정과 근거를 시각적으로 제공함으로써, 실제 비즈니스와 정책의 의사결정 과정을 보다 투명하고 신뢰할 수 있게 지원한다. 다섯째, 클라우드나 고성능 PC 없이도 누구나 현장에서 쉽게 사용할 수 있는 머신러닝 및 딥러닝 기반 비즈니스 애널리틱스 프레임워크를 제시하였다. 누구나 손쉽게 데이터를 학습하고 활용할 수 있도록 설계되었기에 실용성을 극대화하는

연구다.

이후의 내용은 방법론을 통해 17개의 공공빅데이터를 병합하는 과정, 혼잡도 예측을 위해 사용된 머신러닝과 딥러닝 알고리즘의 소개, 그리고 예측 결과와 결론을 제시하며 마무리된다.

2. Methods

2.1. *Collecting and Merging Data*

한국은 정보화 시대의 도래와 더불어 행정의 투명성 제고와 사회적 활용 효율성을 강화하기 위해 공공데이터 개방 정책을 적극 추진해왔다. 이 과정에서 국민들의 정보 접근성이 크게 향상되었고, 데이터 기반 정책 수립은 물론 민간 차원에서 창의적인 활용이 가능해졌다. 1990년대 초반부터 정부는 공공정보의 디지털화를 시도하며 일부 데이터를 개방하기 시작했으며, 2013년 제정된 *「공공데이터의 이용 및 제공에 관한 법률」*은 공공데이터 활용을 제도적으로 뒷받침하는 중요한 전환점이 되었다 ([참고자료](#)). 현재는 중앙정부뿐 아니라 지방자치단체 차원에서도 다양한 공공데이터가 생산·개방되고 있으며, 향후 데이터의 양과 활용 범위는 더욱 확장될 것으로 전망된다. 이러한 맥락에서 공공데이터의 융합적 활용은 데이터 기반 의사결정(data-driven decision making)을 고도화하는 필수 전략으로 자리매김하고 있다.

본 연구는 기존 선행연구에서 제한적으로 활용되던 데이터 범위를 넘어, 공공데이터 융합을 극대화함으로써 지하철 혼잡도 예측 성능을 크게 개선하고자 한다. 구체적으로, "공공데이터 포털, 서울 열린 데이터 광장, 통계청 등" 총 8개의 주요 포털로부터 수집한 17개 데이터베이스(DB)를 통합하였다. <표 1>에 융합한 공공 데이터베이스의 출처와 이름부터 기간, 단위, 추출변수, 그리고 종속변수 활용에 대한 상세 설명을 제시하였다. 이 데이터베이스에는 지하철 역사 정보와 이용객 수뿐 아니라, 자동차 등록 대수, 주유소 판매가격과 같은 교통 관련 지표, 그리고 경기선행지수·실업률·기준금리·환율·소비자물가·인구밀도 등 거시경제 및 사회적 변수가 포함되어 있다. 다시 말해, 본 연구는 지하철 혼잡도를 단일 교통 데이터에 한정하지 않고, 교통·경제·사회적 요인을 포괄적으로 통합한 공공데이터 융합 모델을 구축하였다.

데이터의 시계열적 일관성을 확보하기 위해, 각 데이터베이스가 제공하는 기간(2015~2024년) 중 공통적으로 추출 가능한 구간인 2021~2023년을 분석에 활용하였다. 또한 단위 차이를 극복하기 위해 모든 데이터를 일 단위로 정규화·통합하였다. 이러한 접근은 현존하는 대한민국의 공공데이터를 최대한 활용하여 지하철 혼잡도의 예측 정확도뿐 아니라 정책적 해석 가능성을 동시에 제고한다는 점에서 기존 연구와 차별된다.

표 1. 설명가능한 지하철 혼잡도 예측 성능 개선을 위해 활용한 공공 빅데이터 정보 및 변수활용 방향

데이터출처	데이터베이스(DB)명	기간	단위	추출변수	종속변수활용
공공데이터포털	한국천문연구원_특일 정보	전체	일	연도, 월, 일, 요일, 공휴일, 대체공휴일	-
	서울교통공사_1_8호선 역별 시간대별 승객유형별 승하차인원	2015.01-2024.07	일	일별 승하차인원수, 우대권인원수, 청소년인원수	승강장혼잡도
	서울교통공사_지하철혼잡도정보	2019-2024	분기	분기별 05 시~23 시 상하선 혼잡도(정원대비 승차인원)	-
	서울교통공사_월별 환승유입인원	2019-2023	연	월별 환승유입인원수	-
	서울교통공사_역사운영 현황	2024.06.30 기준	연	역사운영현황(호선, 역명, 면적, 흥수, 승강장유형, 출입구수, 환승노선)	승강장혼잡도
	서울교통공사_수송순위	2019-2023	연	연도별 수송인원수	상위 30 개역 필터링 및 종속변수 결측치 존재역 제외
서울열린데이터광장	서울시 지하철역 정보 검색 (역명)	2024.1.14 기준	연	역명 대응 역코드	-
	서울시 역코드로 지하철역별 열차 시간표 정보 검색	2024.1.14 기준	연	역코드 대응 열차시간표, 시간대별 상하행 평균운행간격, 상하행 운행횟수	승강장혼잡도 2(면적 대비 승강장대기인원)
기상자료개발포털	종관기상관측(ASOS)	전체	시간	일시, 평균기온, 일강수량, 평균풍속, 평균상대습도, 일최심적설	-
Opinet	국내유가통계 > 주유소 > 평균판매가격	전체	일	일별 보통휘발유, 자동차용경유 가격	-
국토교통부 (시도별 자료)	국토교통부 > 자동차 등록 현황	전체	연	연도별 자동차 등록대수	-

통계청「산업활동 동향」	기획재정부 > 경기종합지수	전체	월	월별 동행지수 순환변동치, 선행지수 순환변동치	-
통계청「경제활동 인구조사」	기획재정부 > 취업자 수/실업률 추이	전체	월	월별 실업률, 청년실업률	-
통계청「장래인구 추계 시도편 : 2022-2052」	통계청 > 지역별 인구 및 인구밀도	전체	연	연도별 서울/수도권 인구, 인구밀도	-
한국은행 경제통계시스템(e cos) -> 1.3 금리	금융위원회 > 시장금리 추이	전체	월	월별 기준금리	-
Investing.com	USD/KRW - 미국 달러 원	전체	일	일별 환율	-
	한국 소비자물가지수(CPI)	전체	월	월별 소비자물가지수, CPI 대비 지하철요금	-

본 연구는 서울 지하철 혼잡도를 정밀하게 예측하기 위해, 총 17개의 공공 데이터베이스(DB)를 융합하여 분석에 활용하였다. 그 중 핵심 기준 데이터는 "서울교통공사_1~8호선 역별 일별 시간대별 승객유형별 승·하차 인원"으로, 최근 3년간 분기별 데이터를 수집하였다. 이 데이터는 교통카드 기록을 기반으로, 서울시 1~9호선 전 구간의 일자별·시간대별 승·하차 현황을 제공하며, 호선·역명·승차여부·승객유형별 인원을 포함한다. 변수명 불일치를 보정하였고, 일부 분기에서 발생한 46만 건 이상의 비정상 결측치는 제거하였다. 또한 승·하차 인원 합계 및 차이를 계산하여 파생 변수를 생성하였고, 청소년·우대권 이용객 수를 구분하여 포함하였다.

보조 데이터로 활용한 "서울교통공사_지하철혼잡도정보"는 30분 단위 혼잡도를 연도별로 제공한다. 일부 시간대에서 36~1133건의 결측치가 발견되었으나, 전체 표본 대비 미미하여 제거하였다. 이를 통해 상·하행선별 혼잡도 지표를 구성하였으며, "날짜"를 공통 키(key)로 기준 데이터와 병합(Left Join)하였다. "서울교통공사_월별 환승유입인원" 데이터는 연도별 구조 차이를 보였으나 행 단위로 변환하여 통일하였고, 소수의 결측치는 제거하였다. 결과적으로, 일자·호선·역명·환승 유입인원으로 구성된 데이터를 앞선 자료와 병합하였다.

또한 "서울교통공사_역사운영 현황"을 활용하여 역별 면적·층수·승강장 유형·출입구 수·환승노선 현황을 추가하였다. 환승노선이 없는 경우는 결측치를 0으로 처리하였다. 이 과정에서 승강장 면적을 기준으로 승·하차 인원 대비 면적당 이용밀도 변수를 산출하여, 본 연구의 종속변수로 활용하였다. 아울러 지하철 요금은 성인 기준으로

2023년 10월 1일을 기점으로 1,250원에서 1,400원으로 인상된 변화를 반영하였다.

비교 설명력을 높이기 위해 교통·경제·기상 데이터를 추가적으로 통합하였다. "종관기상관측(ASOS)" 자료에서 평균 기온, 강수량, 풍속, 상대습도, 적설량을 추출하였으며, [Investing.com](#)에서 확보한 "원·달러 환율"과 "한국 소비자물가지수(CPI)"를 결합하였다. CPI 대비 지하철 요금이라는 파생 변수를 추가해 요금 부담을 상대적으로 측정하였다. "오피넷"의 휘발유와 경우 별 주유소 판매가격, "통계청"의 실업률·청년실업률 및 경기종합지수(동행·선행지수), 인구 및 인구밀도, "국토교통부"의 자동차 등록대수, "한국은행"의 월별 기준금리 자료 또한 기준 데이터와 병합하였다.

마지막으로, "서울열린데이터광장"의 지하철역 시간표 정보를 활용하여 주요 30개 역의 평일·토요일·공휴일별 상·하행 운행횟수 및 평균 운행 간격을 산출하였다. 모든 데이터베이스는 "날짜"를 공통 키로 병합하였으며, 이를 바탕으로 연·월·일·요일·주차 변수와 *holidayskr* 라이브러리를 통해 생성한 공휴일 더미 변수를 포함하였다. 이렇게 구축된 융합 데이터셋은 시계열적 일관성과 변수 간 이질성을 보정한 형태로, 지하철 혼잡도 예측과 설명가능한 인공지능(XAI) 적용을 위한 기반 자료로 활용되었다. 이해도를 높이기 위해 전체적인 공공빅데이터 융합 데이터베이스 생성 과정을 제시한다 (**오류! 참조 원본을 찾을 수 없습니다.**).



dbdiagram.io

그림 1. 빅데이터 융합DB 생성을 위한 데이터베이스 구조 및 결합과정

2.2. Data Preprocessing and Understanding

본 연구에서 종속변수는 승강장 혼잡도로 정의되며, 이는 승하차 인원을 승강장 면적으로 나눈 값으로 산출되었다. 전체 약 164만 개의 관측치에 기반한 평균 혼잡도는 45.91이며, 최소 0.01에서 최대 502.33까지 광범위하게 분포한다. **오류! 참조 원본을 찾을 수 없습니다.**의 좌측 패널에서 노선별 분포를 비교하면, 서울 지하철에서 가장

많은 역을 포함하는 2호선의 혼잡도 변동 폭이 다른 노선에 비해 압도적으로 넓게 나타난다. 이어서 3호선과 4호선은 2호선의 절반 수준에서 분포하며, 특정 노선의 구조적 특성과 이용 패턴 차이를 반영한다. **오류! 참조 원본을 찾을 수 없습니다.**의 우측 패널은 역별 혼잡도를 내림차순 정렬한 것으로, 강남역이 가장 넓은 혼잡도 범위를 보였으며, 홍대입구·잠실·신림·서울대입구 등 환승역과 2호선 주요 거점들이 상위를 차지하였다.

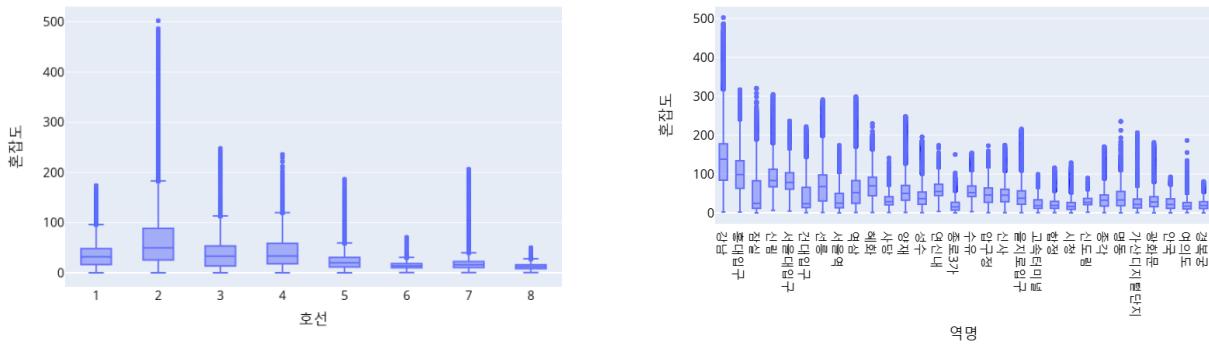


그림 2. 종속변수인 승강장혼잡도의 호선별 그리고 역별 분포 비교

독립변수는 총 113개로 구성되며, 교통·경제·사회적 요인을 포괄하는 공공데이터 융합 변수들이 포함되어 있다. 이들은 승강장 혼잡도를 예측하기 위한 주요 요인으로 활용되었다. 그러나 정책적으로 중요한 의사결정은 혼잡도가 낮은 시점이 아니라 혼잡이 극도로 심화되는 시점에서 필요하다. 이에 따라 본 연구는 수송 인원수가 많은 상위 30개 역을 분석 대상으로 선정하였다. 이 과정에서 결측치가 존재하는 역은 제외하여 최종 분석의 신뢰성을 확보하였다.

전체 데이터셋은 약 164만 개의 샘플을 포함하며, 종속변수와 독립변수를 합쳐 총 114개의 변수로 구성되었다. 노선별로 살펴보면, 분석에 포함된 역은 1호선부터 8호선까지이며, 2호선이 14개 역으로 전체 샘플의 37.01%를 차지하였다. 그 뒤를 이어 3호선은 8개 역으로 20.19%를 차지하였으며, 반면 6호선과 8호선은 각각 단 1개 역만을 포함하였다. 이러한 분포는 데이터셋이 특정 노선에 상대적으로 집중되어 있음을 보여주며, 동시에 대규모 환승 및 중심 상권과 연결된 노선에서 혼잡 예측이 정책적으로 더욱 중요함을 시사한다.

표 2. 통합된 공공 빅데이터 수송인원수 상위 30개 지하철 호선별 샘플수, 비율, 역명 정보

호선	샘플수	비율	역명
2	610,278	37.01%	홍대입구, 시청, 신림, 을지로입구, 선릉, 역삼, 성수, 사당, 강남, 서울대입구, 신도림, 건대입구, 잠실, 합정
3	332,878	20.19%	연신내, 신사, 종로 3 가, 양재, 압구정, 경복궁, 고속터미널, 안국
4	208,050	12.62%	수유, 혜화, 명동, 사당, 서울역
1	166,436	10.09%	종로 3 가, 서울역, 종각, 시청
7	124,830	7.57%	가산디지털단지, 고속터미널, 건대입구
5	124,828	7.57%	광화문, 여의도, 종로 3 가
6	41,610	2.52%	합정
8	40,106	2.43%	잠실

본 연구에서는 생성된 융합 빅데이터를 기반으로 종속변수(승강장 혼잡도)를 예측하기 위해 총 113개의 독립변수를 활용하였다. 그러나 알고리즘이 학습할 수 있도록 하기 위해서는 데이터의 **전처리(preprocessing)** 과정이 필수적이다. 먼저 분석 목적과 무관한 변수를 제거하였으며, 문자형 응답값은 더미 변수(dummy variables) 또는 수치형 변수로 변환하였다. 이어서 변수 간 범위 차이로 인한 왜곡을 최소화하기 위해 모든 수치형 변수는 **스케일링(scaling)** 과정을 거쳐 일정한 범위로 정규화하였다.

데이터 분할은 **학습(Training) 세트와 검증(Validation) 세트**로 구분하였으며, 각 역당 평균 약 5.49만 개의 샘플을 확보하였다. 전체 데이터는 약 8:2의 비율로 무작위(Random sampling) 분할하였다. 비록 연구 데이터가 시계열적 속성을 지니고 있으나, 지하철 혼잡도는 직전 시점의 흐름보다는 **출근·퇴근 시간대, 요일, 공휴일 여부** 등 특정 상황적 패턴에 더 큰 영향을 받는다. 또한 과거 3년간의 동일 시간대 샘플이 충분히 확보되어 있어, 전통적 시계열 알고리즘을 직접 적용하기보다는 이러한 시간적 특성을 **파생 변수(feature engineering)** 형태로 반영하는 것이 적절하다. 이에 따라 데이터는 시간 순서를 보존하지 않고 무작위로 분할하였으며, 대신 연·월·주·요일·공휴일 여부 등의 변수를 추가하여 시계열적 맥락을 반영하였다.

모델 학습 단계에서는 학습 세트를 사용하여 알고리즘을 훈련시켰으며, 평균제곱오차(Mean Squared Error, MSE)를 최소화하는 방향으로 하이퍼파라미터를 최적화하였다. 검증 세트는 학습 과정에서 배제된 데이터를 기반으로, 과거와 유사한

상황이 미래에도 반복될 수 있다는 가정 하에 모델의 일반화 성능을 평가하는 데 활용되었다.

이러한 접근은 단순히 통계적 정확도에 그치지 않고, 실제 정책 현장에서의 활용 가능성을 높이는 데 목적이 있다. 다시 말해, 지하철 혼잡도를 설명하는 모델은 단순히 “예측력이 높은가”를 넘어, 과거 및 유사한 상황에서 승객이 실제로 경험할 혼잡도를 정확히 재현해야 하며, 그 결과는 정책 담당자와 시민 모두에게 의미 있는 해석을 제공해야 한다. 이를 통해 본 연구는 **설명력과 신뢰성을 동시에 갖춘 혼잡도 예측 모델**을 제시한다.

2.3. Statistical Algorithm

본 연구는 지하철 혼잡도 예측의 **baseline 모델**로 선형회귀(linear regression)를 적용하였다. 선형회귀는 가장 전통적이고 직관적인 *statistical learning* 알고리즘 중 하나로, 독립변수와 종속변수 간의 관계를 단순하고 투명하게 해석할 수 있다는 장점 때문에 다양한 분야에서 널리 활용되어 왔다. 본 연구에서는 각 샘플에 해당하는 독립변수들이 가중치와 결합되어 종속변수(혼잡도)를 추정하는 구조를 따른다. 선형회귀의 기본식은 다음과 같이 표현된다:

$$\hat{Y} = f(X_1, X_2, \dots, X_k) = w_0 + w_1X_1 + w_2X_2 + \dots + w_kX_k = XW$$

이를 통해 각 독립변수가 혼잡도에 미치는 상대적 영향력을 확인할 수 있다. 그러나 이 모형은 변수 간 관계를 선형(linear)으로 가정하기 때문에, 실제 지하철 혼잡도의 비선형적·복합적 패턴을 충분히 반영하지 못한다. 따라서 본 연구에서는 선형회귀를 비교·평가기준(baseline)으로 설정하고, 보다 정교한 머신러닝 및 딥러닝 기법들과의 성능을 비교하였다.

2.4. Machine Learning Algorithm: Bagging and Boosting

머신러닝은 인공지능 기법 중 예측 오차를 최소화하기 위해 고안된 방법론으로, 오차를 편향(bias)과 분산(variance)으로 구분하여 접근한다. 이 과정에서 분산을 줄여 예측의 안정성을 높이는 배깅(Bagging)과, 편향을 줄여 성능을 개선하는 부스팅(Boosting)이 대표적인 전략으로 활용된다 (Chen and Guestrin 2016). 두 방법론

모두 분류(classification)와 회귀(regression) 문제에 적용 가능한 **지도학습(supervised learning)** 기법이다.

배깅의 대표적 알고리즘인 랜덤 포레스트(Random Forest, RF)는 복수의 샘플을 부트스트래핑(sampling with replacement)하여 병렬적으로 학습한 뒤, 투표(voting)나 평균을 통해 최종 예측값을 산출한다. RF는 계산 속도가 빠르고 학습 데이터에 대한 예측력이 우수하며, 노이즈 데이터에 대해서도 강건한(robust) 성능을 보인다 (Michie, Spiegelhalter, and Taylor 1999). 또한 변수 중요도(feature importance)를 제공하여 정책적 시사점을 제시할 수 있으나, 각 변수의 영향 방향성(positive/negative)을 명확히 설명하지 못한다는 한계가 존재한다.

부스팅은 개별 모델의 약점을 반복적으로 보완하며 성능을 고도화하는 방식으로, Gradient Boosting Machine (GBM)을 기반으로 다양한 확장 알고리즘들이 개발되었다. 대표적으로 **XGBoost**, **LightGBM**, **CatBoost**는 과적합(overfitting) 방지를 위한 정규화 기능, 메모리 효율성 향상, 연산 속도 개선 등 다양한 기능이 추가되어 있다 (Alsubari et al. 2021). 이러한 알고리즘들은 초기 학습에서 정확도가 낮았던 샘플의 가중치를 점차 강화하면서 재학습을 수행하여, 높은 예측 성능을 달성할 수 있다.

2.5. Deep Learning Algorithm: MLP, CNN

딥러닝은 인간 두뇌의 신경망 구조에서 영감을 받아 개발된 인공지능 방법론으로, 다층 구조(layered architecture)를 통해 데이터 내 잠재된 **복잡한 비선형 패턴과 변수 간 상호작용**을 학습하는 데 강점을 지닌다. 그 기초 구조는 다층 퍼셉트론(Multi-Layer Perceptron, MLP)으로, 입력층-은닉층-출력층으로 구성되어 있으며, 회귀와 분류 문제 모두에 적용 가능한 범용 근사기(universal approximator) 역할을 한다.

이러한 기본 구조는 데이터 유형에 맞게 확장되었으며, 특히 **합성곱 신경망(Convolutional Neural Network, CNN)**은 이미지 인식 분야에서 특화되었지만, 본질적으로는 입력 변수들의 **국소적(local) 상관 구조**를 추출·요약하는 데 강점이 있다. CNN은 합성곱(convolution), 풀링(pooling), 전결합(fully connected) 계층을 거치면서 저차원(low-level) 특성에서 점차 고차원(high-level) 특성으로 변수 표현을 변환한다 (Shustanov and Yakimov 2017). 본 연구에서는 이러한 과정을 통해 **지하철 혼잡도에 영향을 미치는 다양한 변수들의 상호작용 패턴**을 학습하였으며, 최종 출력층에서

혼잡도 값을 예측하는 회귀 문제로 접근하였다.

2.6. Model Explainability: SHapley Additive exPlanations (SHAP)

머신러닝 및 딥러닝은 뛰어난 예측 성능에도 불구하고, 결과가 산출된 이유를 설명하기 어려운 블랙박스(**black box**) 문제를 안고 있다. 일부 알고리즘이 변수 중요도(feature importance)를 제공하더라도, 이는 영향의 크기만을 보여줄 뿐 방향성(positive/negative)을 설명하지 못한다. 이러한 한계를 보완하기 위해 제안된 방법이 **SHapley Additive exPlanations (SHAP)**이다. SHAP은 **LIME(Local Interpretable Model-Agnostic Explanations)**과 **게임이론 기반 Shapley** 값을 결합한 기법으로, 모델 예측에 기여하는 변수의 정도와 방향성을 동시에 해석할 수 있다 (Molnar 2023).

구체적으로, LIME은 특정 입력 값을 교란(perturbation)시켰을 때 예측값의 변화를 가중치로 환산하며, Shapley 값은 협력게임이론(coalitional game theory)에 기반하여 각 변수가 기여하는 평균적 효과를 계산한다. 계산량이 방대할 경우에는 무작위 샘플링(random sampling)을 활용하여 효율성을 확보한다. 이를 통해 SHAP은 복잡한 딥러닝 및 머신러닝 모델의 내부 작동 원리를 **정량적이고 직관적인 설명**으로 변환한다.

전통적인 통계 알고리즘은 직관적 해석이 가능하지만 낮은 예측 성능과 불안정성을 보인다. 반대로 인공지능 알고리즘은 높은 성능을 보이지만 설명력이 부족하다. 따라서 본 연구는 **고성능 딥러닝 예측 모델과 SHAP 기반 설명 가능성**을 결합하여, 단순히 "정확한 예측"을 넘어 "왜 그러한 예측이 나왔는가"에 대한 신뢰할 수 있는 해석을 제공한다. 이는 정책 입안자와 도시 교통 관리자가 실제로 활용 가능한 **실질적·현실적 의사결정 도구**로서 의미를 갖는다.

2.7. Evaluation Metrics

지하철 혼잡도 예측 모델의 성능을 다각도로 검증하기 위해 본 연구는 6가지 회귀 성능 지표(regression metrics)를 사용하였다. 이는 단일 지표에 의존하는 경우 발생할 수 있는 편향을 최소화하고, 모델의 예측력을 보다 종합적으로 평가하기 위함이다. 선정된 지표는 RMSE(Root Mean Squared Error), MSPE(Mean Squared Percentage Error), MAE(Mean Absolute Error), MAPE(Mean Absolute Percentage Error), MedAE(Median

Absolute Error), MedAPE(Median Absolute Percentage Error)이다. 각 지표의 정의는 다음과 같다.

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

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

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

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{|y_i|}$$

$$MedAE = median(|\hat{y}_1 - y_1|, |\hat{y}_2 - y_2|, \dots, |\hat{y}_n - y_n|)$$

$$MedAPE = median\left(\left|\frac{\hat{y}_1 - y_1}{y_1}\right|, \left|\frac{\hat{y}_2 - y_2}{y_2}\right|, \dots, \left|\frac{\hat{y}_n - y_n}{y_n}\right|\right)$$

모든 지표는 실제값(y_i)과 예측값(\hat{y}_i)의 차이를 기반으로 산출되며, 수치가 낮을수록 모델의 예측 성능이 우수함을 의미한다. 특히, 본 연구와 같이 **대규모·이질적** 변수로 구성된 **빅데이터 환경**에서는 특정 지표에만 의존하기보다 다양한 지표를 병행하여 검증하는 것이 필수적이다. 따라서 본 연구는 여섯 가지 지표를 종합적으로 활용함으로써, 제안된 모델이 단순히 높은 정확도를 보이는 것에 그치지 않고, 실제 **정책적·실무적 활용에서 신뢰 가능한 수준의 설명력과 안정성을 갖추었는지 평가하였다.**

3. Results

이번 섹션에서는 전처리 완료된 종속변수와 독립변수를 사용하여 모델에 데이터를 학습 지하철 혼잡도의 예측 결과를 제시한다. 실제 지하철 혼잡도 예측 성능이 높아야 모델이 설명하는 기여도 또는 관련성과 같은 설명 신뢰성이 확보될 수 있을 것이다. 선형회귀분석과 같은 전통적인 알고리즘을 사용하는 경우 지하철 혼잡도 예측을 위한

변수 기여는 쉽게 알수가 있지만 실제 그러한 변수들의 설명력으로 혼잡도를 예측할 경우 성능이 낮은 경향이 있다. 전통적인 사회과학에서 설명력에 집중하느라 그러한 설명이 미래에 얼마나 비즈니스적으로 신뢰할 수 있는 결과인지는 경시하는 경향이 있다. 따라서 머신러닝과 딥러닝으로 forecasting performance 결과를 우선적으로 확인하여 모델링의 성능을 확보하고 다음으로 블랙박스와 같은 알고리즘에 SHAP 알고리즘을 적용하여 신뢰할 수 있는 변수들의 설명력을 제공한다. 데이터 준비, 전처리, 모델링, 성능 검증 등의 모든 데이터분석 프로세스는 python 3.9.20 버전을 사용하였다. 그리고 Machine Learning 알고리즘들은 sklearn 1.5.2 버전의 라이브러리를 그리고 Deep Learning 알고리즘은 tensorflow 2.17.0 버전의 라이브러리를 사용하였다. 그리고 알고리즘의 예측 성능을 설명하기 위해 사용한 SHAP는 0.46.0 버전의 라이브러리를 사용하였다.

3.1. Prediction Performance

이번 section에서는 지하철 혼잡도의 모델링 성능을 확인하기 위해 머신러닝 및 딥러닝 알고리즘의 미래 예측 성능을 소개한다. 예측문제를 해결하기 위한 전통적인 알고리즘인 Linear Regression, 대표적인 머신러닝 알고리즘인 Random Forest, XGBoost, LightGBM, 그리고 CatBoost를 포함하여 총 5가지 머신러닝 알고리즘을 사용하였다. 그리고 대표적인 딥러닝 알고리즘인 MLP and CNN 총 2가지 알고리즘을 사용하였다. 전체 데이터 80%의 training set을 사용하여 모델링하고 hyperparameters를 조정하였다. 그리고 20% test set을 사용하여 미래 예측 성능을 확인함으로써 모델의 성능을 확보한다. 알고리즘을 성능을 공평하게 비교하기 위해 가급적 동일한 parameter values를 사용하였다 (표 1).

표 3. 지하철 혼잡도 모델링에 사용한 알고리즘의 하이퍼파라미터

Algorithm	Parameter	Value
Logistic Regression	Intercept	TRUE
	Max Iteration	100
	Random State	0
Random Forest XGBoost LightGBM CatBoost	Samples	100
	Criterion	Gini
	Minimum Spitting Samples	2
	Random State	0
MLP	TensorFlow Random Seed	0
	Number of Hidden Layers	2

CNN	Unit Numbers for Hidden Layers	(100, 100)
	Activation for Hidden Layers	Relu
	Activation for Output Layer	identity
	Loss	Mean Squared Error
	Optimizer	Adam
	Epochs	500
	TensorFlow Random Seed	0
	Number of Convolution Layers	1
	Unit Numbers for Convolution Layers	64
	Number of Dense Layers	2
	Kernel Size	2
	Stride	1
	Padding	Valid
	Unit Numbers for Dense Layers	64
	Activation for Hidden Layers	Relu
	Activation for Output Layer	Linear
	Loss	Mean Squared Error
	Learning Rate	0.001
	Optimizer	Adam
	Batch Size	32
	Epochs	10

알고리즘에 데이터를 학습시킨 후, 지하철 혼잡도의 Test 예측 성능을 검증한다. 각 역마다의 모든 알고리즘의 검증지표 순위를 추정한 후 30개역의 순위 평균치(표 4)와 각 알고리즈다 30개역의 검증지표 평균값(표 5)를 통해 보다 일반화된 알고리즘의 설명이 가능하다. 표 4를 보면 RMSE의 경우 MLP 알고리즘이 30개역 평균 순위가 1.31로 나타난다. 즉, 대부분의 경우에서 RMSE 지표는 MLP 알고리즘의 지하철 혼잡도 예측 순위가 1등으로 나타남을 의미한다. 그런데 RMSE를 제외한 나머지 5개 검증지표들에서는 Random Forest 알고리즘이 30개역 평균 순위가 1.15와 1.00으로 나타났다. 따라서 나머지 검증지표들에선 Random Forest 알고리즘이 거의 항상 1등을 했다는 것이다. 따라서 6개의 검증지표들 중 5개의 검증지표에서 상위순위를 차지한 알고리즘은 Random Forest다. 그리고 다음 순위를 차지한 알고리즘은 MLP다. 실제 알고리즘별로 30개역의 검증지표 수치를 평균해 보았더니(표 5), 1순위를 차지한 Random Forest 알고리즘이 평균치 기준인 MSPE에서 0.02%, MAPE에서 0.50%를 차지할만큼 1% 미만의 오차를 보이며 매우 정확하게 지하철 혼잡도를 예측하는 것으로 나타났다. 2순위를 차지한 MLP의 경우도 MSPE와 MAPE가 각각 0.34%와 1.83%를 나타낼 정도로 높은 정확도를 보이고 있다. 반면 전통적인 선형회귀분석 알고리즘은

순위에서도 최하위이며 실제 예측 오류는 MSPE가 144% 만큼 매우 부정확한 예측임을 알 수 있다. 따라서 이러한 정확도를 보이는 모델링에서 변수들의 기여도 또는 설명력을 신뢰하는 것도 어렵다고 판단된다. 따라서 지하철 혼잡도 예측에 사용한 공공 빅데이터 융합DB의 독립변수들 패턴은 단순히 선형회귀분석에서 잘 학습될 만큼 선형(Linear)패턴이 아니고 훨씬 복잡한 비선형(Non-linear) 패턴이 많이 포함되어 있음을 의미하고, 고성능의 PC 성능을 요구하는 고성능 딥러닝 알고리즘이 아니더라도 공공 빅데이터를 융합할 경우 매우 정교하고 높은 예측 모델링을 구축할 수 있음을 의미한다.

표 4. 수송인원수 상위 30개 역의 미래 예측 성능 검증지표 순위 평균

Algorithm	RMSE	MSPE	MAE	MAPE	MedAE	MedAPE
Linear Regression	6.95	6.95	7.00	6.95	7.00	6.95
Random Forest	1.74	1.15	1.00	1.00	1.00	1.00
XGBoost	4.23	4.38	4.23	4.33	4.36	4.31
LightGBM	4.82	4.69	4.85	4.77	4.82	4.85
CatBoost	6.00	5.90	5.97	5.90	5.87	5.90
MLP	1.31	1.85	2.00	2.00	2.00	2.00
CNN	2.95	3.03	3.00	3.00	3.00	3.00

표 5. 알고리즘별 수송인원수 상위 30개 역의 미래 예측 성능 검증지표들 중 퍼센트 오차 평균

Ranking	MSPE	MAPE	MedAPE
Random Forest	0.02%	0.50%	17.47%
MLP	0.34%	1.83%	89.16%
CNN	0.90%	4.30%	234.97%
XGBoost	3.68%	9.86%	675.02%
LightGBM	5.89%	10.82%	732.84%
CatBoost	20.73%	16.69%	1052.98%
Linear Regression	144.05%	48.20%	2865.35%

일반화된 성능 이외에 실제 얼마나 지하철역의 혼잡도를 잘 예측하는지 확인하기 위해 미래 월별 혼잡도와 특정일 시간대별 혼잡도 예측 성능을 확인한다. 우선 **오류! 참조 원본을 찾을 수 없습니다.**의 오른쪽에 혼잡도 1위로 분포하고 있는 강남역과 가장 하위에 분포하고 있는 경복궁역을 대상으로 하였다. 가장 최근 개통된 신분당선이 환승되기도 하며 국내의 가장 많은 유동인구가 있기에 강남역을 대상으로 하는 것은

중요하며, 경복궁역 역시 대한민국의 역사적, 상징적, 그리고 문화적으로도 관광객이 많이 이용하는 곳이기 때문에 국내외 이용자들에겐 매우 관심이 있는 혼잡도 정보가 될 것이다. 두 역 모두 Random Forest 알고리즘이 가장 높은 성능을 보였고 비교를 위해 두 번째로 높은 성능인 MLP 알고리즘과 함께 월별(그림 3) 그리고 시간대별(그림 4) 혼잡도 예측을 시각화하였다. 월별 미래 예측 성능에서도 강남역과 경복궁역 모두 Random Forest 알고리즘이 매우 실제값과 유사한 혼잡도를 예측하는 것을 확인할 수 있고, 시간대별 예측 성능은 더욱 정확하게 예측하는 모습을 보인다. 따라서 개별 지하철역에서도 미래 혼잡도를 예측하는데 매우 효과적인 방법임을 확인할 수 있다. 이제 매우 신뢰성 높은 모델을 구축하였으니 비즈니스 의사결정에 직접적으로 활용될 수 있는 신뢰성 있는 변수들의 기여 또는 설명력을 제공할 수 있다.

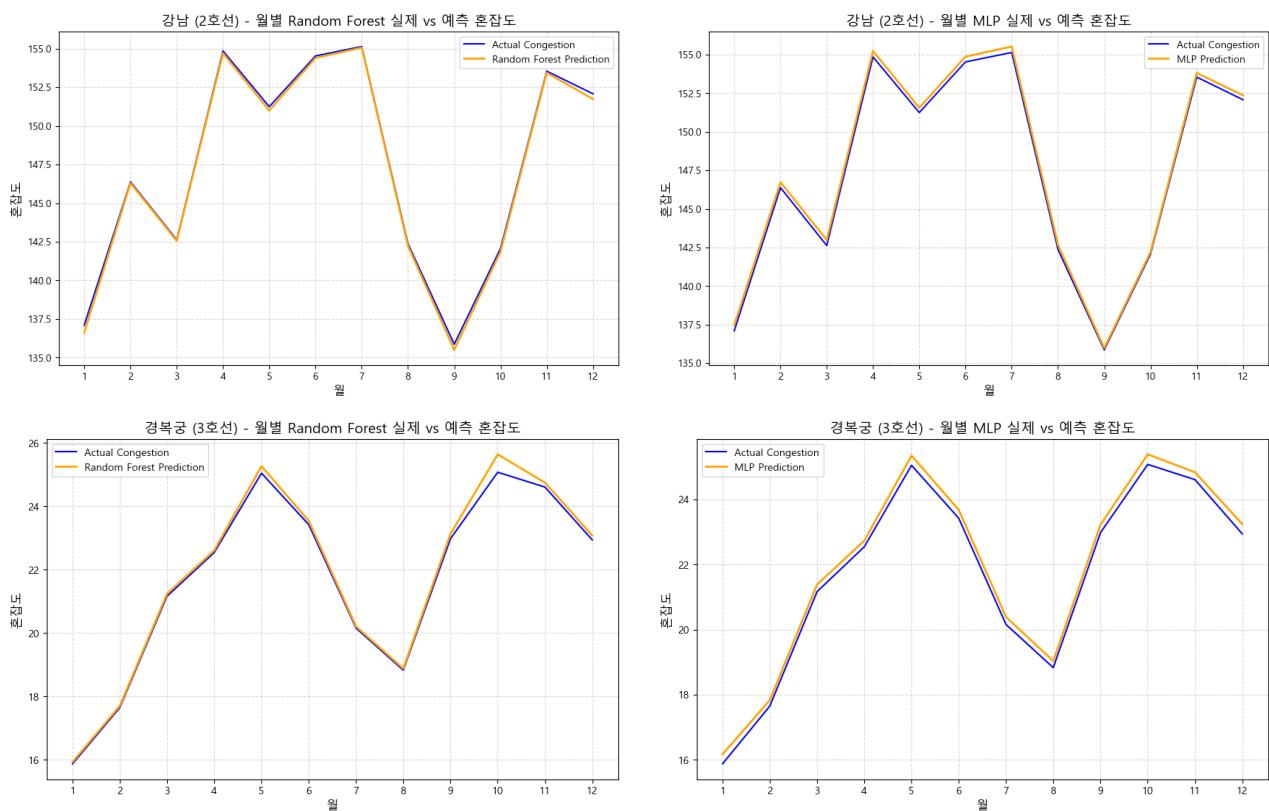


그림 3. 강남역과 경복궁역의 월별 Random Forest(1등)와 MLP(2등) 알고리즘 미래 예측 시각화

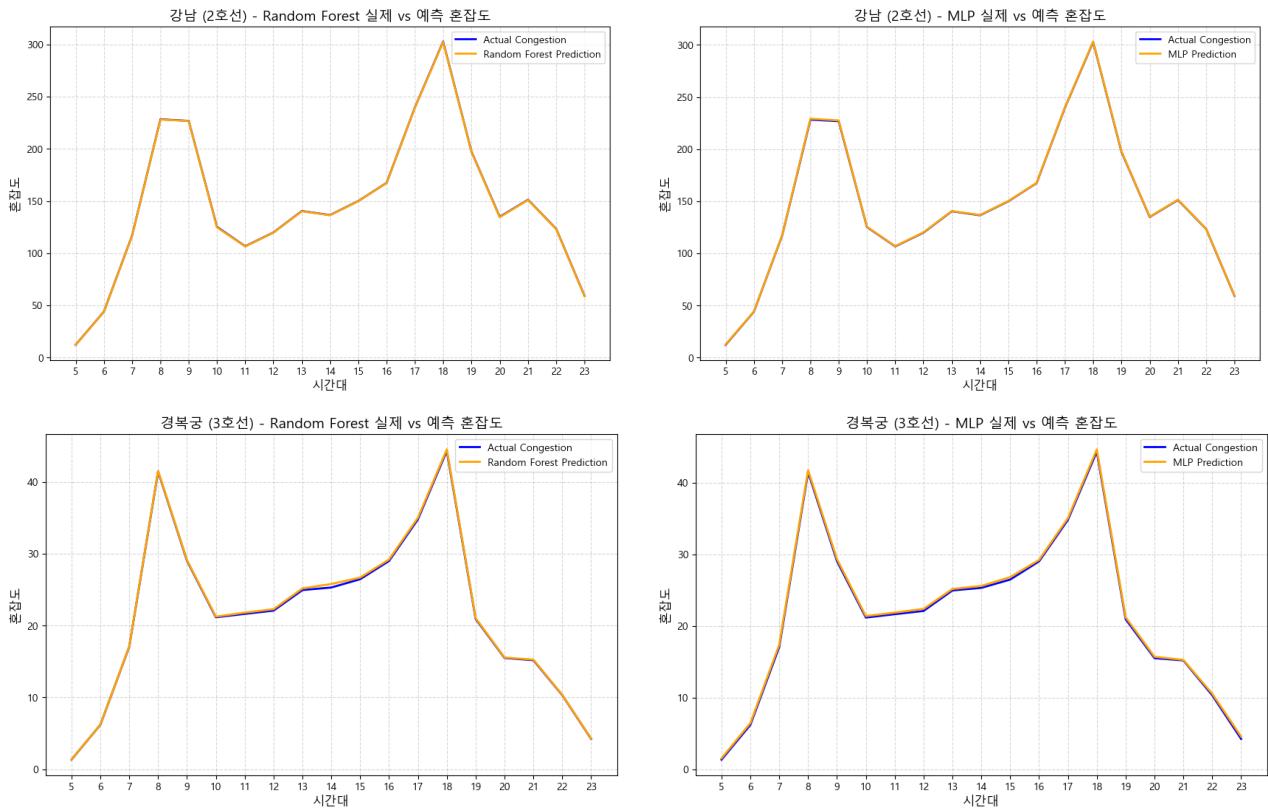


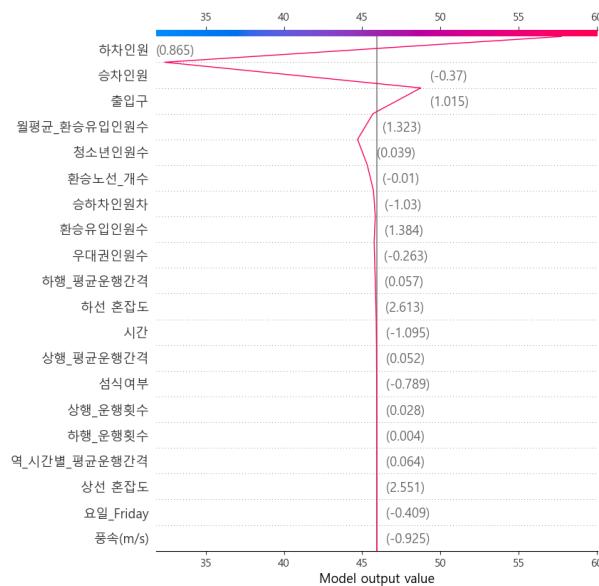
그림 4. 강남역과 경복궁역의 시간대별 Random Forest(1등)와 MLP(2등) 알고리즘 미래 예측 시각화

3.2. Explainability of 지하철 혼잡도 Prediction

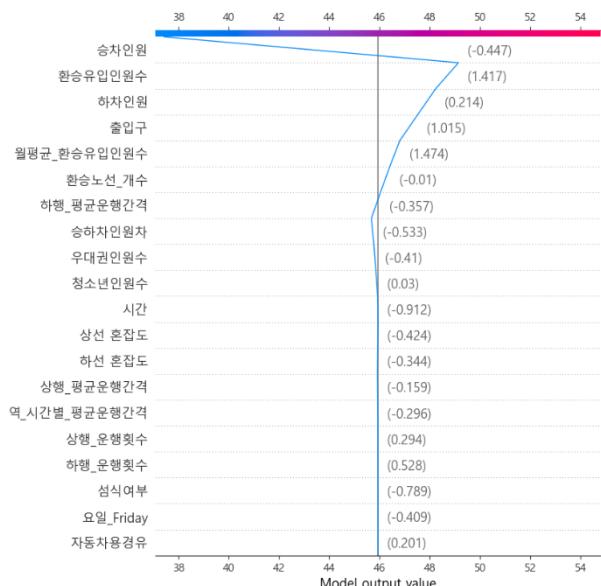
머신러닝 그리고 딥러닝과 같은 인공지능 알고리즘은 복잡한 비선형 패턴과 변수들의 모든 상호작용 조합까지도 학습을 하여 모델을 만들기 때문에 전통적인 linear regression과 달리 예측에 기여하는 독립변수들의 기여도를 단순하게 파악하기 어렵다. 물론 머신러닝 알고리즘은 딥러닝과 달리 데이터의 학습 과정에서 예측 성능을 높이는데 기여한 feature importance를 제시하긴 하지만 순위들을 단순하게 누적할 뿐 positive or negative effect와 같은 방향성조차 제시하지 못한다. 하지만 SHAP explainer를 사용하면 지하철 혼잡도의 높은 미래 예측 성능이 왜 그러한지 독립변수들의 상대적 기여도로 평가할 수 있다. SHAP는 독립변수들의 수치들만 수집되면 실시간으로 지하철의 혼잡도가 얼마가 될 것인지 위 section에서 파악한 매우 높은 성능으로 예측해 낼 수 있다. 뿐만 아니라 독립변수들이 어떠한 방향으로 기여해서 그러한 높은 미래 예측 성능을 달성하였는지 설명해주기 때문에, 실제 지하철 교통 비즈니스에 활용되기도 용이하고 신뢰성도 높아 의사결정에 용이하게 활용될 수 있다.

그리고 이러한 실시간 혼잡도처럼 누적된 혼잡도 예측과 설명력들을 모두 결합하여 일반화된 정책으로 활용할 수 있는 지하철 혼잡도의 신뢰성 높은 설명력을 시각적으로 표현하기 때문에 이해하기 쉽다.

우선 실시간 지하철 혼잡도에 대한 설명력을 decision plot으로 시각화 할 수 있다. 가장 높은 성능을 보이는 Random Forest 알고리즘을 기준으로, 그림 5는 강남역(상단)과 경복궁역(하단)의 미래 특정일을 임의로 선택하여 혼잡도를 예측하고 어떻게 혼잡도 수치가 나오게 되었는지 변수들의 기여를 내림차순으로 보여준다. 예를 들어 강남역의 혼잡도가 높을 때 약 58점 정도의 수치가 예측되었는데 “하차인원, 출입구, 월평균_환승유입인원수” 등은 예측 혼잡도를 높이는 변수들이고 “승차인원, 청소년인원수, 환승노선_개수” 등은 예측 혼잡도를 낮추는 변수들이다. 그외 나머지 변수들은 혼잡도에 크게 기여하지 못하는 모습이다. 하지만 강남역 혼잡도가 낮을 때는 “승차인원”이 주된 혼잡도를 낮추는 변수이며 “환승유입인원수, 하차인원, 출입구, 월평균_환승유입인원수, 환승노선_개수, 하행_평균운행간격” 등은 혼잡도를 높이는 변수들로 나타난다. 미래 특정일과 상황에 따라 변수들이 혼잡도를 높이거나 낮추는 영향정도와 방향은 얼마든지 달라질 수 있으며 실시간으로 높은 정확성으로 혼잡도를 예측할 뿐만 아니라 어떻게 그러한 수치가 나오게 되는지 의사결정 근거로 활용하기에도 매우 용이하다.



(강남역 혼잡도 높을 때)



(강남역 혼잡도 낮을 때)

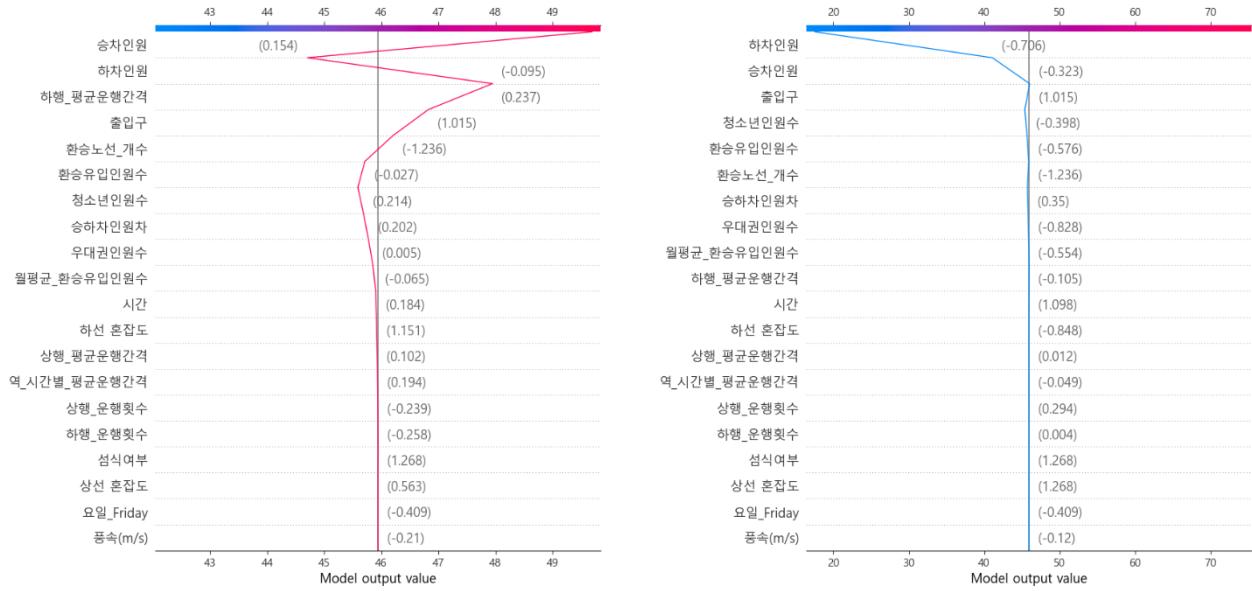


그림 5. 미래 특정 일 대상 강남역과 경복궁역의 혼잡도 예측과 변수영향 설명력 기반 의사결정

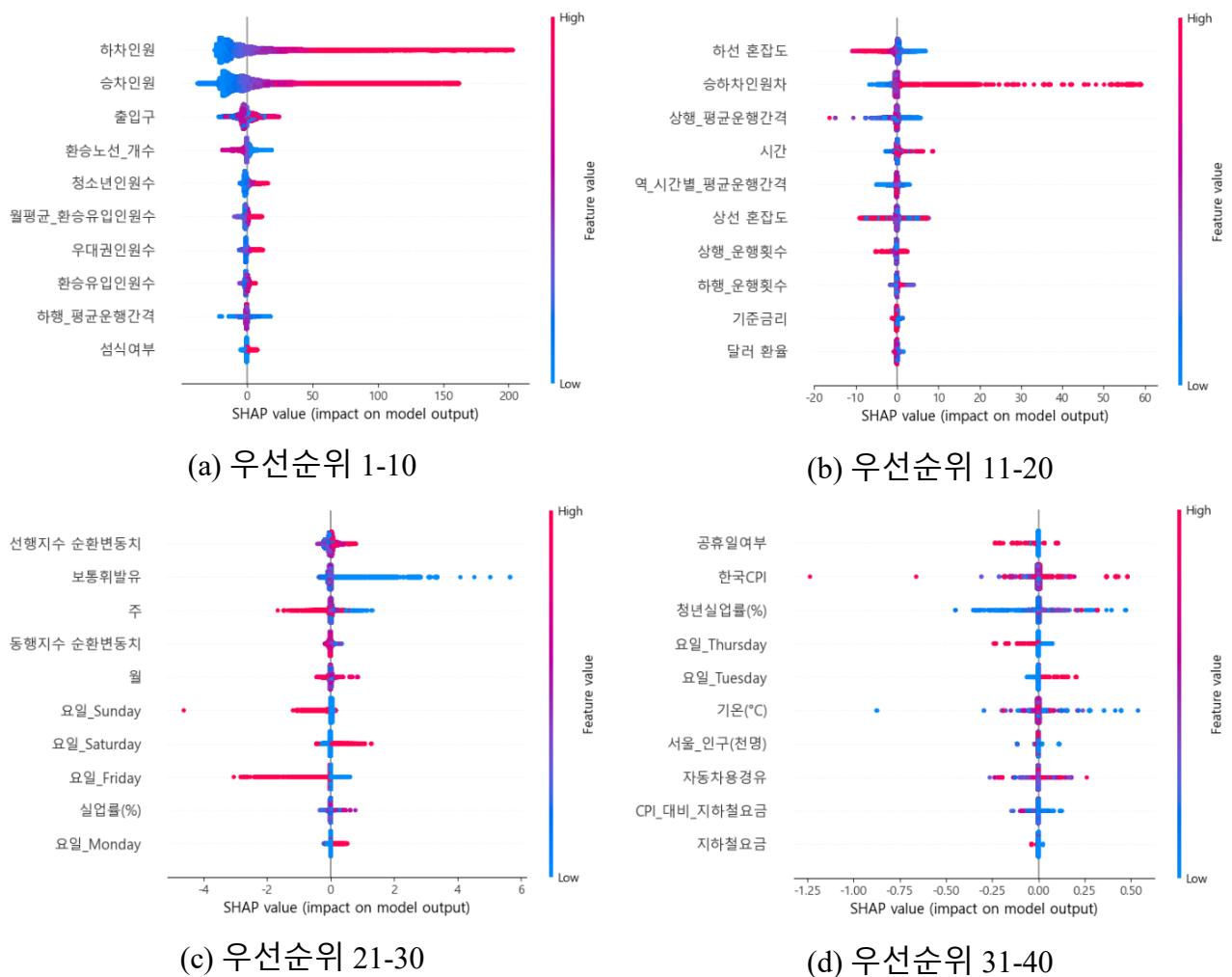
특정일 또는 시간대의 지하철 혼잡도에 대한 변수들의 기여를 확인하였다. 각 변수가 지하철 혼잡도에 일반화된 positive or negative를 보장하지도 않는다. 그러한 이유로 특정 변수가 지하철 혼잡도에 어떠한 방향성을 가지는지 일반화하는 것은 다소 위험할 수 있다. 정책적 근거를 위해 일시적으로 사용할 수는 있겠지만 정책 입안자들 또는 시민들에게 납득될만한 정보로 단정하기 어렵기 때문이다. 그렇다고 정책적 근거로 사용하지 말자는 것은 아니며, 높은 정확성의 실시간 혼잡도의 예측 설명력을 확인할 수 있었으니 이를 누적하여 표현한다면 충분히 일반화된 인사이트로 재표현 할 수 있을 것이다. 단, positive or negative 와 같은 이분법적인 해석은 매우 조심해야 함을 위의 실시간 예측 결과를 통해 다시 한번 강조한다. 실제 독립변수와 종속변수의 관계가 이분법적 설명으로 명확할 정도로 선형적 관계이지도 않다. 만약 선형적 관계라면 선형회귀분석의 결과가 성능이 좋아야 하는데 전혀 그렇지도 않았기 때문이다. 따라서 본 연구에서는 매우 명확한 관계성이 나타나는 변수에 대해서만 조심스럽게 해석을 제안하면서 머신러닝 딥러닝 기반의 설명가능한 인공지능의 효과적인 활용법을 가이드하는데 집중할 것이다.

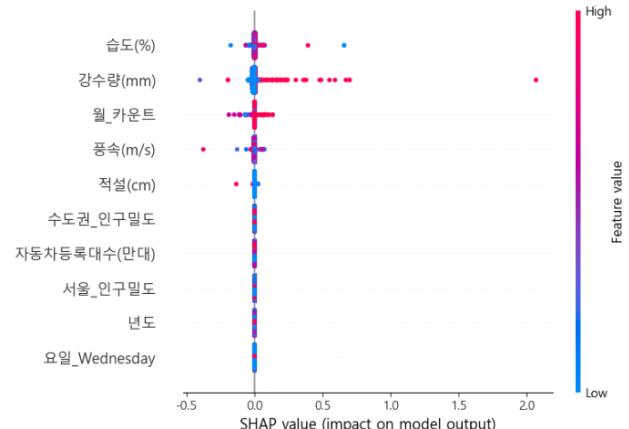
각 샘플마다의 지하철 혼잡도 예측 설명력을 누적하여 일반화된 설명력을 확인할 것이다. 단, 본 연구에 활용한 샘플 수 기준의 일반화이기 때문에 샘플이 늘어나거나

줄어들 경우 변경될 수도 있음을 인지해야 한다. 아마도 데이터의 샘플이 매우 커진다면 향후 그 일반화 설명력을 더욱 신뢰할 수 있을 것임은 분명하다. 그림 6는 Random Forest 알고리즘으로 모든 미래 데이터의 지하철 혼잡도 예측을 위한 feature impacts를 시각화 한 것이다. 우선 세로축은 왼쪽과 오른쪽으로 2개가 있다. (왼쪽)세로축은 종속변수에 영향을 주는 feature importance 를 내림차순으로 정렬한 것이기 때문에 높은 곳에 위치한 변수들은 낮은 곳에 위치한 변수들보다 지하철 혼잡도 예측에 더욱 많이 기여한다. 즉, "하차인원, 승차인원, 환승노선_개수, 출입구, 하행_평균운행간격, 청소년인원수, 승하차인원차" 등의 순서로 변수들의 기여순위가 분석된다. 그리고 (오른쪽)세로축은 변수들의 값이 낮은 경우는 파란색 계열로 그리고 높은 경우는 빨간색 계열로 표시하여 변수들의 값에 대응되는 지하철 혼잡도의 예측값의 변화를 표시하였다. 마지막으로 가로축에 바로 지하철 혼잡도의 예측값을 반영하였는데, 예측에 대한 각 독립변수들의 average marginal contribution이다. 따라서 SHAP 값이 0보다 작으면 negative contribution을 0보다 크면 positive contribution을 의미한다. 각 시점마다 변수의 수치는 다양할 수 있고 각 수치별 지하철 혼잡도 or feature impact 방향도 다양할 수 있기 때문에, 각 변수들의 값과 대응되는 SHAP output 위치를 모두 누적하여 표현하면 각 변수의 값의 변화에 따른 지하철 혼잡도 기여 방향이 시각화 되는 것이다.

예를 들어 "하차인원과 승차인원" 변수는 값이 작을때(파란색계열) negative contribution에 많이 분포되어 있고 값이 커지면(빨간색계열) positive contribution에 많이 분포가 되어 있다. 따라서 해당 변수의 값이 높을수록 지하철 혼잡도가 높아지도록 영향을 준다고 해석할 수 있다. 반대로 "환승노선_개수" 변수는 값이 작은 수치에서 커질수록 negative contribution에 많이 분포되기 때문에 지하철 혼잡도를 낮추는데 영향을 준다고 해석할 수 있다. 이를 정리하여 미래 예측 성능이 가장 높은 지하철 혼잡도에 positive correlation을 보이는 가장 중요한 특징은 "하차인원, 승차인원, 청소년인원수, 월평균_환승유입인원수, 우대권인원수, 환승유입인원수, 섬식여부, 승하차인원차, 시간, 선행지수 순환변동치, 요일_Saturday, 실업률, 요일_Monday, 요일_Tuesday" 등이며, negative correlation을 보이는 특징은 "환승노선_개수, 하선 혼잡도, 상행_평균운행간격, 기준금리, 달러환율, 주, 동행지수 순환변동치, 요일_Sunday, 요일_Friday, 요일_Thursday, 지하철요금" 등으로 해석할 수 있다. 단순히 지하철 관련 변수들 뿐만 아니라 각 시간정보에 따른 혼잡도 기여, 그리고 경기상황이나

경제지표들의 영향력도 확인할 수 있으며 이것인 기존 연구들에서 발견하기 어려웠던 결과다. 총 17개의 DB 결합과 설명가능한 인공지능을 통해 확인할 수 있었다. 나머지 변수들도 유사한 해석을 할 수 있지만 값의 변화와 contribution 방향이 명확한 정/역방향인 경우 그리고 그 contribution 변화의 폭이 넓은 경우만 정책적인 근거로 활용하길 추천한다. 그렇지 않은 애매한거나 좁은 범위의 변화를 억지로 일반화 하는 것은 왜곡된 해석을 줄 수 있으니 지양하는 것이 좋다.





(e) 우선순위 41-50

그림 6. 지하철 혼잡도 예측에 대한 변수들의 기여 설명력 기반 의사결정

4. Conclusions

최근 전세계적으로 확산되었던 코로나 바이러스 역시 사람들이 밀집되는 폐쇄적 대중교통에서 강한 전염성을 보여 지하철 이용이 크게 감소하였기 때문에, 의학적으로나 서비스적으로 혼잡도 이슈는 다양한 방향으로 확산되기 쉽다. 글로벌 시장조사업체 스태티스타의 분석에 따르면 전세계에서 대중교통 이용률이 가장 높은 나라는 대한민국이라고 한다. 사회적 이슈들을 개선하기 위해서는 다수가 이용하는 대표적 교통수단인 지하철의 혼잡도를 사전에 예측하는 것은 서비스의 효율성과 그리고 시민들의 편의성과 안정성 측면에서 중요한 의미를 갖는다.

본 연구에서는 시민들이 객차 내외부 혼잡도를 반영하기 위해 승강장 내 혼잡도를 예측하는 것을 목적으로 한다. 총 17개의 데이터베이스를 융합하고 설명가능한 인공지능을 활용하여, MSPE 기준 99% 이상의 높은 혼잡도 예측이 가능하였다. 그리고 예측 결정 과정과 근거를 시각적으로 제공함으로써 실시간으로 비즈니스에 활용할 수 있을 뿐 아니라 의사결정 과정을 보다 투명하고 신뢰할 수 있도록 기여하는 비즈니스 애널리틱스를 제시하였다. 그 결과 승하차인원, 청소년이나 노인층의 탑승 여부 뿐만 아니라 경기상황들이 반영되는 선행지수나 실업률 등이 지하철 혼잡도를 높이는데 기여함을 알 수 있었다. 반면 여러 노선이 겹쳐서 환승이 가능하거나 상행선 평균운행간격, 그리고 경제상황이 반영된 기준금리, 달러환율, 동행지수, 지하철 요금 등은 혼잡도를 낮추는데 기여할 수 있음이 확인되었다. 이러한 결과들은 다수의 연구에서 발견된 것과 유사한 부분도 있지만 이전 연구에서 발견되지 않은

경제지표와의 관련성을 발견하는데 기여하였다.

본 연구는 고성능 PC가 아니더라도 설명 가능한 인공지능을 활용하여 지하철 혼잡도를 실시간으로 확인할 수 있는데 도움을 줄 수 있기 때문에, 서비스 측면에서 시민들에게 빠르고 효과적인 대응책을 마련하는데 도움이 될 것이다. 나아가 향후 수도권 인근 신도시의 개발이나 다양한 형태의 교통수단들의 수요분배 효과를 파악하는데도 큰 도움이 될 것이다.

Data Availability

The data can be provided on request.

Declaration of Competing Interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Jaeheung Park: Conceptualization, Software, Validation, Data Curation, Writing Original Draft, Preparation. **Kyungwon Kim:** Methodology, Preparation, Software, Writing Review and Editing, Visualization, Supervision.

References

- Adadi, and Berrada. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI), *IEEE Access, Access, IEEE*, 6: 52138-52138.
- Alsubari 등., (2021). Development of Integrated Neural Network Model for Identification of Fake Reviews in E-Commerce Using Multidomain Datasets, *Applied Bionics and Biomechanics*, 2021: 5522574.
- Cha 등., (2021). Comparison of random forest and gradient boosting machine models for predicting demolition waste based on small datasets and categorical variables, *International journal of environmental research and public health*, 18: 8530.
- Chen, and Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. San Francisco, California, USA: Association for Computing Machinery.
- Fleck. 2023. 'Fare Play: Who's Riding Public Transit?', Statista, Accessed 01.07. <https://www.statista.com/chart/30882/share-of-commuters-that-take-public-transport/>.
- Michie 등., (1999). Machine Learning, Neural and Statistical Classification, *Technometrics*, 37.
- Molnar. (2023). *Interpretable machine learning. A Guide for Making Black Box Models Explainable*, Lulu. com.
- Office. (2023). *Seoul Transportation in 2022*, edited by Futuristic Advanced Transportation Division
- Shustanov, and Yakimov. (2017). CNN Design for Real-Time Traffic Sign Recognition, *Procedia Engineering*, 201: 718-725.
- Viswavandya 등., (2021). ANALYSIS AND COMPARISON OF MACHINE LEARNING APPROACHES FOR TRANSMISSION LINE FAULT PREDICTION IN POWER SYSTEMS, *Journal of Research in Engineering and Applied Sciences*, 6: 24-31.
- 권효승 등., (2020). GRU 를 이용한 서울시 지하철 구간별 혼잡도 예측 모델 연구, *한국 ITS 학회/학술대회*: 185-190.
- 김규진 등., (2024). 딥러닝 분석을 이용한 지하철역 혼잡도 변화의 동적 분석 및 예측 모델에 관한 연구, *한국통신학회 학술대회/논문집*, 2024: 1192-1193.
- 김승준. (2016). *서울시 지하철의 혼잡비용 산정과 정책적 활용방안*, 1-22,
- 김재익. (2013). 아침 첨두시간대 지하철 이용수요의 결정요인에 관한 연구 - 대구 지하철 역세권 토지이용을 중심으로, *교통연구*, 20: 15-25.
- 김진수. (2016). 빅데이터 분석을 이용한 지하철 혼잡도 예측 및 추천시스템, *디지털융복합연구*, 14: 289-295.
- 성현곤. (2017). 서울시 대중교통 수단별 월별 이용수요의 변동에 영향을 미치는 요인 분석, *국토계획*, 52: 81-96.
- 이경재 등., (2020). 심층신경망 모형을 활용한 대중교통 이용자의 환승시간 추정에 관한 연구, *한국ITS 학회/논문지*, 19: 32-43.

- 이상준. (2020). 교통카드 자료를 활용한 지하철 혼잡도 개선 연구 : Early Bird 정책대안을 중심으로, *한국IT서비스학회지*, 19: 125-138.
- 이상준 등,. (2021). 승객 승하차 이동 및 대기행태를 고려한 도시철도 승강장 점유면적 추정방안, *대한교통학회지*, 39: 721-735.
- 이선하 등,. (2015). 철도 이용객 정보제공 효과평가 방법론 연구 -승강장의 혼잡상황을 고려한 Gate Metering 사례 연구 중심으로, *한국ITS학회 논문지*, 14: 50-62.
- 이정훈, and 정현영. (2018). 분위 회귀를 활용한 기상조건이 대중교통 수단별 통행량에 미치는 영향에 대한 연구, *국토계획*, 53: 95-106.
- 이호, and 최진경. (2015). 대중교통카드 자료를 활용한 도시철도 승강장 혼잡도 추정 알고리즘 개발, *한국철도학회논문집*, 18: 270-277.
- 장진영 등,. (2023). 머신러닝 기반 2호선 출퇴근 시간대 지하철 역사 내 혼잡도 예측, *문화기술의 융합*, 9: 145-150.
- 천예은 등,. (2021). 설명 가능한 AI 기술을 활용한 신용평가 모형에 대한 연구, *한국데이터정보과학회지*, 32: 283-295.