
딥러닝 기반 KTX 수송수요 예측 연구

- 자문보고서 -

2023. 7.

제 출 문

한국철도공사 연구원장 귀하

이 보고서를 「딥러닝 기반 KTX 수송수요 예측 연구」의 자문 최종보고서로 제출합니다.

2023. 7.

자 문 가 : 성균관대학교 소프트웨어융합대학 교수 한진영

제 1 장 서 론

1.1 연구의 배경 및 목적

1.2 연구의 범위 및 수행체계

제1장 서론

1.1 연구의 배경 및 목적

1.1.1 연구배경

- KTX 수송 수요 예측의 고도화를 위하여 최근 좋은 성능을 보이고 있는 딥러닝 기반 시계열 모형의 개발 작업이 필요함
 - KTX 수송 수요에 대한 다양한 형태의 예측을 시도하고 딥러닝 기반 시계열 분석 방법론을 적용하고 고속열차의 노선별 수송수요 예측을 수행

1.1.2 연구목적

- 여객사업본부에서 활용되는 2024년도 수송계획 및 열차운영계획 수립에 필요한 단기수송수요예측을 위한 딥러닝 기반 시계열 모델을 통한 수송수요 예측 방법 구축을 목적으로 함
- 기존의 수송수요 분석 및 단기 예측 방법을 고도화하기 위하여 고속열차 대상 노선별 및 정차역별 수송수요 예측을 위한 딥러닝 기반 시계열 모델 활용 결과 분석을 진행함
 - 다양한 변수로부터 발생하는 영향을 고려하기 위하여 딥러닝 기반 시계열 모델을 구축하고 이를 통해 KTX 노선별 및 정차역별 수송수요 예측을 수행함

1.2 연구의 범위 및 수행체계

1.2.1 연구의 범위

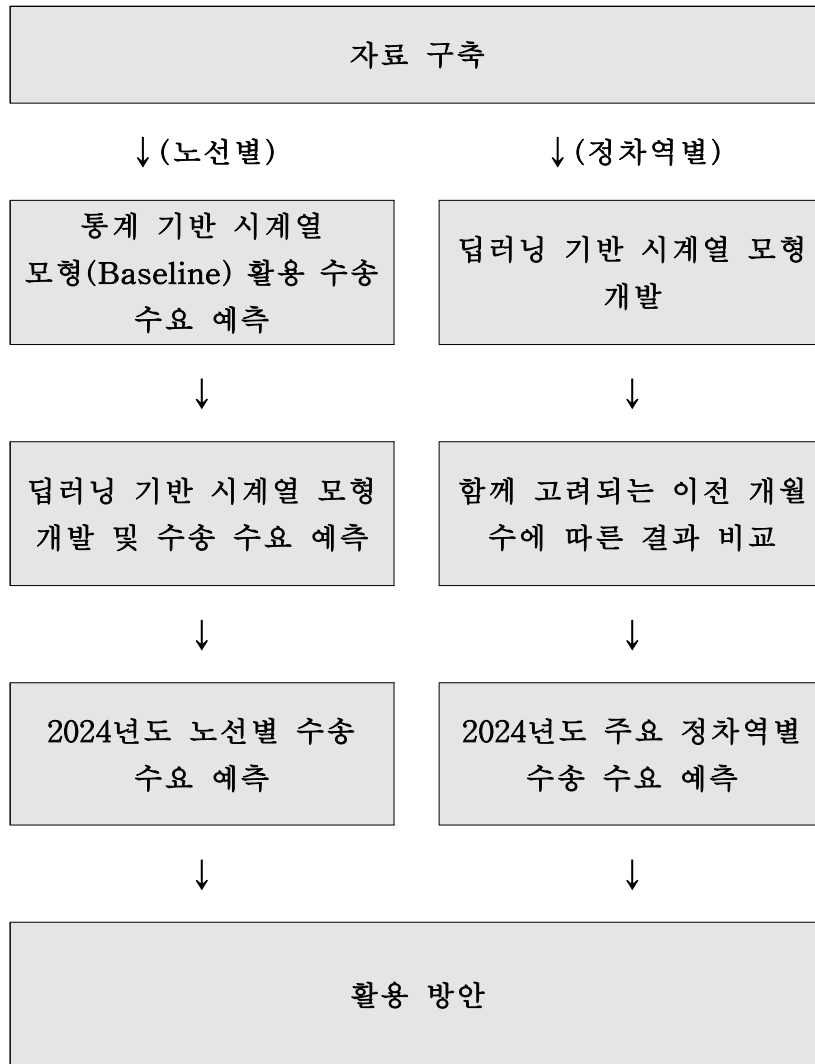
- 본 연구는 KTX 수송 수요 예측을 위한 딥러닝(Deep Learning) 기반 시계열 모형을 개발하기 위해 다음 사항을 전제함
 - 고속열차(즉, KTX, KTX-산천, KTX-호남)가 운행되는 5개 노선(즉, 경부선, 호남선, 전라선, 경전선, 동해선)의 수송수요에 차이가 있으므로 이들 차이를 반영할 수 있도록 구분하여 개발함
 - 보다 정확한 수송수요 예측을 위한 공급좌석 수, 승차율, 이용율 등의 추가 변수를 활용하여 분석함
- 2020년부터 2022년 후반까지의 수송수요 시계열 자료는 COVID-19의 영향으로 인해 일

반적인 상황의 추세적 변화와 유사하지 않은 동향을 보임

- 안정적인 시계열 자료 분석을 위해서는 2020년부터 2022년 후반까지의 자료를 활용하지 않는 것이 좋으나, COVID-19가 완화됨에 따른 수송수요 회복 추세를 반영하기 위하여 해당 기간 자료를 포함하는 것이 적절해 보임
 - 다만, 일반적 상황의 KTX 수송수요 추세를 함께 고려하기 위하여 2015년부터 2019년까지 약 4년치 자료와 COVID-19에 의한 수송수요 감소가 회복되고 있는 2022년 후반부터 2023년 4월까지의 자료를 함께 활용함
- KTX 노선별 수송수요 예측 - 통계 기반 시계열 분석 모형(즉, ARIMA, Prophet)과 딥러닝 기반 시계열 모형(즉, LSTM)을 비교 분석함
 - 자료형태 : 노선별 월별 수송실적 (단, 동해선과 구동해선은 통합하여 활용함)
 - 분석기간 : 2015년 1월 ~ 2023년 4월
 - KTX 정차역별 수송수요 예측 - 딥러닝 기반 시계열 모형(즉, LSTM)을 활용하여 분석하였으며, 함께 고려하는 이전 개월 수의 변화에 따른 결과를 관찰함
 - 자료형태 : 노선별 역별 월별 수송실적 (단, 동해선과 구동해선은 통합하여 활용함)
 - 분석기간 : 2015년 1월 ~ 2023년 4월

1.2.2 연구의 수행체계

- 2024 KTX 노선별 수송수요 예측을 위한 딥러닝 기반 시계열모형 개발의 수행체계는 다음과 같음
 - 2015년 01월 ~ 2023년 04월(약 9년)에 해당하는 노선별 월별 수송실적 데이터 구축
 - 통계 기반 시계열 모형 분석
 - 딥러닝 기반 시계열 모형 제안 및 구축
 - 통계 기반 시계열 모형과 딥러닝 기반 시계열 모형 예측결과 비교
 - 2024년 KTX 노선별 수송수요 예측
 - KTX 수송 수요 예측의 적정성 및 활용 방안 검토
- 2024 KTX 정차역별 수송수요 예측을 위한 딥러닝 기반 시계열 모형 개발의 수행체계는 다음과 같음
 - 2015년 01월 ~ 2023년 04월(약 9년)에 해당하는 정차역별 월별 수송실적 데이터 구축
 - 함께 고려되는 이전 개월 수에 따른 따른 정차역별 수송수요 예측결과 분석
 - 2024년 KTX 주요 3개 정차역별 수송수요 예측
 - KTX 수송 수요 예측의 적정성 및 활용 방안 검토



<그림 1-1> 2024년 01월 - 12월 KTX 수송수요 예측을 위한 딥러닝 기반 시계열 모형 개발의 수행체계

제 2 장 수요 예측 방법 정립

2.1 KTX 수송실적 자료 분석 및 전처리

2.2 시계열 분석 개요

2.3 통계 기반 시계열 모형 개요

2.4 딥러닝 기반 시계열 모형 제안

제2장 수요 예측 방법 정립

2.1 KTX 수송실적 자료 분석 및 전처리

2.1.1 KTX 수송실적 자료 분석

- 노선별 수송실적 자료는 5개 노선(즉, 경부선, 경전선, 동해선, 호남선, 전라선)에 대한 월별 자료로 구성되어 있음
 - 아래는 각각 총량/경부선/호남선/전라선/경전선/동해선의 2015년 1월부터 2023년 4월 기간의 월별 일평균 수송 실적 시계열 그림임



<그림 2-1> 월별 총량 수송실적 시계열도



<그림 2-2> 경부선 월별 총량 수송실적 시계열도



<그림 2-5> 경전선 월별 총량 수송실적 시계열도



<그림 2-6> 동해선 월별 총량 수송실적 시계열도



<그림 2-3> 호남선 월별 총량 수송실적 시계열도



<그림 2-4> 전라선 월별 총량 수송실적 시계열도

- <그림 2-1>은 주운행선에 무관하게 모든 수송 실적을 합한 결과를 나타낸 표임
 - <그림 2-2>, <그림 2-3>, <그림 2-4>, <그림 2-5>, <그림 2-6>은 주운행선의 5개 노선 경부선/호남선/전라선/경전선/동해선의 수송실적을 나타낸 표임
 - 시계열들은 코로나가 처음 나타난 2020년 1월을 기점을 급격하게 하락한 모습을 보임
 - 2022년 여름에 거리두기가 점차 완화되면서 이전 수요로 점차 회복하는 모습을 보임
 - 코로나 시기를 제외하면 2022년부터 2023년은 꾸준히 상승 중임
- 주운행선의 경우 노선마다 평균 월별 승차인원수에 차이가 있음을 <표 2-1>에서 확인함
- 경부선은 유일하게 백만 명대의 월별 승차인원수를 보였으며 경부선, 호남선, 경전선,

전라선, 동해선 순서대로 수가 적어짐

주운행선	평균 월별 승차인원수
경부선	2774600
호남선	779822
전라선	468593
경전선	470207
동해선	389331

<표 2-1> 주운행선 평균 월별 승차인원수

- 시계열 자료의 특징 중의 하나는 시간의 흐름에 따라 자료의 시계열적인 성질이 변할 수 있는데, 이를 고려하지 않고 자료를 분석하는 경우 오히려 예측 정확도에 문제가 발생할 소지가 있음
 - 현재 시점과 밀접한 관계가 있을 것으로 보이는 가까운 과거의 자료를 사용하는 것이 효과적인 대안이 될 수 있음
 - 2024년 수송수요 예측을 위해 최근의 자료와 가장 동일한 시계열적인 성질이 있다고 판단되는 2023년 수송실적 자료를 이용함

2.1.2 KTX 수송실적 자료 전처리

- 노선별 데이터를 예측하기 위해 활용한 변수들은 주운행선/운행년도/운행월/운행횟수/공급좌석수/승차율/이용률/월별일수/휴일수/코로나/전월수송량이 있음 (<표 2-2> 참고)
 - **주운행선**: 열차종을 뜻하며 KTX, KTX-산천, KTX-호남에 대한 주요한 5개 노선인 경부선/호남선/전라선/경전선/동해선이 있음
 - **운행년도**: 해당 열차가 운행된 연도를 의미하며 운행월은 월을 의미함
 - **운행횟수**: 해당 노선에 운행한 고속 열차 운행횟수를 의미하며 전체/대수송/임시/확정 4가지 종류로 세부적으로 집계됨
 - **공급좌석수**: 열차 내에 판매할 수 있는 좌석수를 의미하며 전체/대수송/임시/확정 4가지 종류로 세부적으로 집계됨
 - **승차율**: 승객이 타고 간 거리를 열차가 운행한 좌석거리로 나눈 비율을 의미하며 연인 거리에서 좌석거리를 나누어 구함
 - **이용률**: 공급에 비해 실제로 사용된 승객 인원을 의미하며 승차인의 수에 공급좌석 수를 나누어 구함
 - **월별일수**: 한 달에 일의 개수로 각 달의 상이한 일을 참고하기 위해 포함함
 - **휴일수**: 한 달에 휴일 개수를 의미함
 - **코로나**: 코로나 시기에 있는지 여부를 0과 1로 표현한 변수를 의미함. 2020년 2월부터 2022년 4월까지의 코로나 기간으로 1의 값을 갖고 그 외 기간은 0의 값을 가짐
 - **전월수송량**: 그 전월의 수송인원수를 의미함

변수	예시
주운행선	경부선
운행년도	2022
운행월	2
운행횟수	4,035
공급좌석수	3,527,406
승차율	0.65766
이용률	1.03497
월별일수	31
휴일수	10
코로나	0
전월수송량	324573

<표 2-2> 노선별 데이터의 변수 예시

- 부적절한 자료 제거
 - 동해선의 다른 월의 승차인원수 평균이 389331인 것에 비해 동해선 2015년 3월의 승차인원수는 612로 적절하지 않은 값을 집계 (2015.4월 동해선(포항) 개통)
 - 따라서 이상치이자 부적절한 자료로 간주되어 제거함

변수	예시
출발역	서울
운행년도	2022
운행월	2
월별휴일	1
출발열차수	246
월별_승차율	1.93
월별_이용률	3.82
월별일수	31
휴일수	10
코로나	0
전월수송량	324573

<표 2-3> 정차역별 데이터 변수

- 정차역별 데이터를 예측하기 위해 활용한 변수들은 출발역/운행년도/운행월/월별휴일/출발열차수/월별_승차율/월별_이용률/월별일수/휴일수/코로나/전월수송량이 있음 (<표 2-3> 참고)
 - **출발역**: 열차가 출발한 역을 의미함
 - **운행년도**: 해당 열차가 운행된 연도를 의미하며 운행월은 월을 의미함

- **월별휴일:** 설연휴와 추석연휴를 포함하는지 여부를 뜻하는 변수로 설연휴 혹은 추석연휴가 포함된 달이면 1을, 아닌 달은 0으로 처리함
 - **출발열차수:** 해당 월에 역에서 출발한 열차 수를 의미하며 전체/경부선/호남선/전라선/경전선/동해선 총 6가지 세부 조건으로 집계함
 - **월별 승차율:** 승객이 타고 간 거리를 열차가 운행한 좌석거리로 나눈 비율을 의미하며 본래 일별 데이터였던 역별 데이터를 활용해 월별 승차율을 계산함
 - **월별 이용률:** 공급에 비해 실제로 사용된 승객 인원을 의미하며 본래 일별 데이터였던 역별 데이터를 활용해 월별 이용률을 계산함
 - **월별일수:** 한 달에 일의 개수로 각 달의 상이한 일을 참고하기 위해 포함함
 - **휴일수:** 한 달에 휴일 개수를 의미함
 - **코로나:** 코로나 시기에 있는지 여부를 0과 1로 표현한 변수를 의미함. 2020년 2월부터 2022년 4월까지의 코로나 기간으로 1의 값을 갖고 그 외 기간은 0의 값을 가짐
 - **전월수송량:** 그 전월의 수송인원수를 의미함
- 자료가 부족한 출발역 제외
- 정차역별 열차 데이터는 2018년 1월부터 2023년 4월까지 데이터로 총 64개월에 대한 데이터를 다룰 수 있음
 - 이 중 장성/김제/서대구/인천공항T1/인천공항T2/검암 역들은 많게는 44개월, 적게는 1개월의 데이터만 집계되었음을 확인
 - 데이터가 상대적으로 부족한 역들을 제외하고 충분한 데이터가 있는 39개 역을 모델 학습에 활용하였음

2.2 시계열 분석 개요

- 역별/정차역별 수송실적 자료를 분석하여 수송수요를 예측함
 - 역별/정차역별 2024년 수송실적 예측을 위해 각각 역별/정차역별 수송실적을 분석한 후 수송 수요에 대한 예측값을 생성함
 - 이를 위해 노선별 예측에서는 통계 기반 시계열 분석 모형인 ARIMA, Prophet과 함께 딥러닝 기반 시계열 모형 LSTM을 비교 분석함
 - 정차역별 수요 예측에서는 딥러닝 기반 시계열 모형인 LSTM을 활용함
- 노선별 수송실적 자료 분석에 대한 내용을 아래 <표 2-4>에 요약하였으며 자세한 내용은 다음과 같음
 - 통계 기반 시계열 분석 모형인 ARIMA, Prophet과 함께 딥러닝 기반 시계열 모형 LSTM을 비교 분석함
 - 분석에 사용될 노선은 총 5개 주운행선(경부선/호남선/전라선/경전선/동해선)으로 동해선과 구동해선은 동해선으로 통합하여 활용함
 - 예측 위해 사용된 자료 구축 기간은 2015년 1월부터 2023년 4월까지임

- 이 중 자료 분석 및 모델 학습에 사용한 기간은 2015년 1월부터 2022년 4월이며, 분석 검증 기간은 2022년 5월부터 2023년 4월까지임
- 분석 검증 단계에서 사용된 자료를 바탕으로 2024년의 예측값을 생성함

구 분		설 명
분석자료		KTX 노선별 수송실적 예측
분석방법		통계 기반 시계열 분석 모형(즉, ARIMA, Prophet)과 딥러닝 기반 시계열 모형(즉, LSTM)을 비교 분석
분석 기간	경부선	2015년 1월 ~ 2022년 4월
	호남선	2015년 1월 ~ 2022년 4월
	전라선	2015년 1월 ~ 2022년 4월
	경전선	2015년 1월 ~ 2022년 4월
	동해선	2015년 1월 ~ 2022년 4월
검증기간		2022년 5월 ~ 2023년 4월
최종예측기간		2023년 4월 ~ 2024년 12월

<표 2-4> KTX 노선별 수송실적 분석 개요

- 정차역별 시계열 모형 분석에 대한 내용을 아래 <표 2-5>에 요약하였으며, 자세한 내용은 다음과 같음
 - 정차역별 예측은 딥러닝 기반 시계열 모형인 LSTM을 활용하여 분석하였으며 함께 고려하는 이전 개월 수의 변화에 따른 결과를 관찰함
 - 분석에 사용된 역은 총 39개 역이 있음
 - 예측을 위해 사용된 자료 구축 기간은 2018년 1월부터 2023년 4월까지임
 - 이 중 자료 분석 및 모델 학습에 사용한 기간은 2018년 1월부터 2022년 4월이며, 분석 검증 기간은 2022년 5월부터 2023년 4월까지임
 - 분석 검증 단계에서 사용된 자료를 바탕으로 2024년의 예측값을 생성해 KTX 주요 3개 정차역인 서울, 부산, 동대구역의 수송수요를 예측

구 분	설 명
분석 자료	정차역별 월별 시계열 모형 분석
분석 방법	딥러닝 기반 시계열 모형(즉, LSTM)
검증 분석기간	2018년 1월 ~ 2022년 4월
검증기간	2022년 5월 ~ 2023년 4월
최종예측 기간	2024년 1월 ~ 2024년 12월

<표 2-5> 정차역별 KTX 수송수요 시계열 모형 분석 개요

2.3 통계 기반 시계열 모형 개요

2.3.1 ARIMA(Autoregressive Integrated Moving Average) 개요

- 통계 기반 시계열 모형 중 하나인 ARIMA는 자기회귀(Autoregressive)과 이동평균(Moving Average)을 모두 고려하는 모형으로, 시계열의 비정상성을 설명하기 위해 관측치 간의 차분을 사용한다는 특징이 있음
 - 자기회귀(Autoregression): 이전 관측값의 오차항이 이후 관측값에 영향을 주는 모형으로, (식 1-1) 은 가장 기본이 되는 식으로, ϕ 는 자기상관계수, ϵ 는 화이트 노이즈를 의미함.
 - 이 때 시간차는 1 혹은 그 이상이 될 수 있음

$$AR(1): X_t = \phi X_{t-1} + \epsilon_t \quad (\text{식 1-1})$$

- 누적(Integrated): 차분을 이용하는 시계열 모형을 의미함
- 이동평균(Moving Average): 관측값이 이전의 연속적인 오차항의 영향을 받음을 의미함
- (식 1-2) 은 가장 기본적인 $MA(1)$ 모형을 나타낸 것으로, β 는 이동평균계수, ϵ 는 시점 t 에서의 오차항임

$$MA(1): X_t = \epsilon_t - \beta_1 \epsilon_{t-1} \quad (\text{식 1-1})$$

- 통계 기반 시계열 모형인 ARIMA는 다음과 같이 설명할 수 있음
 - 시계열 자료 Z_t 가 계절주기가 s 인 계절형 ARIMA 모형을 따른다고 할 때 (식 1-3) 형태의 수식으로 표현함

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D Z_t = \delta + \theta_q \Theta_Q(B^s) \epsilon_t, \quad \epsilon_t \sim iid \ N(0, \sigma_\epsilon^2) \quad (\text{식 1-3})$$

- 여기서, B 는 후행연산자(backshift operator)이며 δ 는 평균 상수임
- $\phi(B)$ 와 $\theta(B)$ 는 각각 비계절형 AR(autoregressive) 연산자 및 MA(moving average) 연산자이며 시계열 자료의 비계절형 자기상관성을 설명하기 위해 사용함
- $\Phi(B^s)$ 와 $\Theta(B^s)$ 는 각각 계절형 AR 연산자 및 MA 연산자이며 시계열 자료의 계절형 자기상관성을 설명하기 위해 사용함
- d 와 D 는 각각 비계절형 및 계절형 차분 차수로서 확률적 추세와 계절성이 있는 경우에 $d=1$, $D=1$ 또는 $D=1$ 을 주로 사용함
- p 와 P 는 각각 비계절형 및 계절형 AR차수, q 와 Q 는 각각 비계절형 및 계절형 MA차수임
- Z_t 가 차분 계절형 시계열 모형을 따를 때, $Z_t \sim ARIMA(p, d, q)(P, D, Q)s$ 라 표현함

2.3.2 Prophet 개요

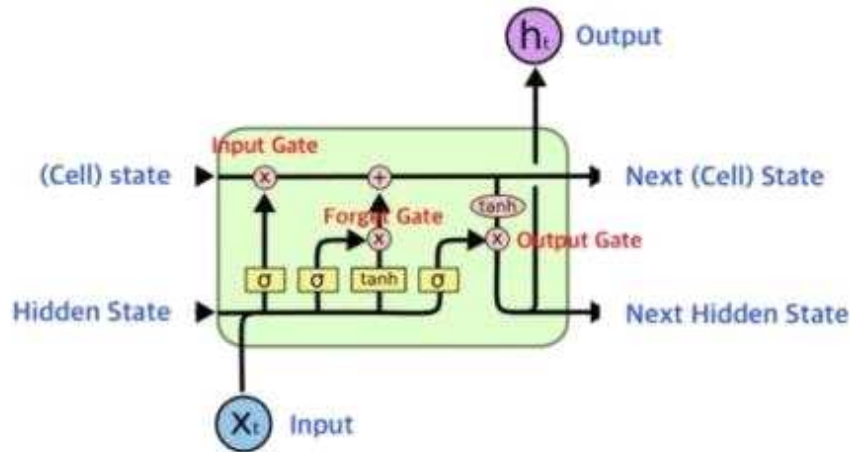
- Prophet은 2017년에 Facebook(현 Meta)에서 시계열 데이터를 다루기 위하여 개발한 라이브러리로, 기존의 통계 기반의 시계열 모형인 ARIMA, ETS, TBATS 등의 한계점 및 복잡성을 단순화한 모형임
- Prophet은 (식 1-4)와 같이 여러 요소를 더하여 사용하는 additive 모형임

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \quad (\text{식 1-4})$$

- $g(t)$ 는 시간의 변화에 따라 비주기성을 띠고 있는 변화를 모델링하며, 선형 혹은 S자 모형으로 모델링함
- $s(t)$ 는 주, 월, 연 단위의 주기성을 보이는 변화를 감지하고 모델링함
- $h(t)$ 는 휴일, 휴가 등 비정기적인 사건과 같은 변화를 모델링함
- ϵ_t 는 모델링이 되지 않는 특이한 변화를 모델링함
- Prophet의 주요 기능은 아래와 같이 서술 가능함
 - 포화도 예측: 성장 또는 감소의 최대치 혹은 최소치를 설정하여 포화도를 예측할 수 있어 시장의 크기 또는 규제 등으로 인한 제약 가능성을 반영하여 예측할 수 있음
 - 경향 변경점: 상승 또는 감소의 경향이 변경되는 지점을 판단 및 구분 가능
 - 계절 변동성/휴일 효과/회귀요소: 계절에 따른 변화나 휴일 또는 주말과 같은 효과를 포함하여 예측 진행 가능함
 - 곱셈 계절성: 계절성은 기본적으로 덧셈(additive)의 형태로 적용되나, 곱셈(multiplication)의 형태로 적용하면 시간의 흐름에 따라 진폭이 커지는 방식으로 예측할 수 있음

2.4 딥러닝 기반 시계열 모형 제안

- 본 연구에서 제안하는 딥러닝 기반 시계열 모형은 딥러닝 시계열 예측 모형 LSTM(Long Short-Term Memory)를 활용하여 설계됨
- LSTM은 RNN(Recurrent Neural Network)의 한 종류로, RNN의 장기 의존성 문제를 해결한 딥러닝 모형이며 직전 데이터뿐만 아니라 보다 거시적으로 과거 데이터를 고려하여 최종적으로 미래 데이터를 예측함



<그림 2-7> Long Short-Term Memory의 cell 개요도

- LSTM은 여러 개의 cell로 구성되어 있으며, 각 cell 내 기능은 <그림 2-7>과 같음
- Cell State: 입력으로 들어온 정보가 바뀌지 않고 그대로 흐르게 하여 상태가 오래 경과하더라도 기울기를 전파시키는 역할을 함
- Forget gate: 과거의 정보 유지 여부를 결정하는 과정으로, (식 1-5)와 같이 시그모이드 함수를 통해 신호를 생성하여 그 값이 0이면 정보를 버리고 1이면 유지하게 됨

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (\text{식 1-5})$$

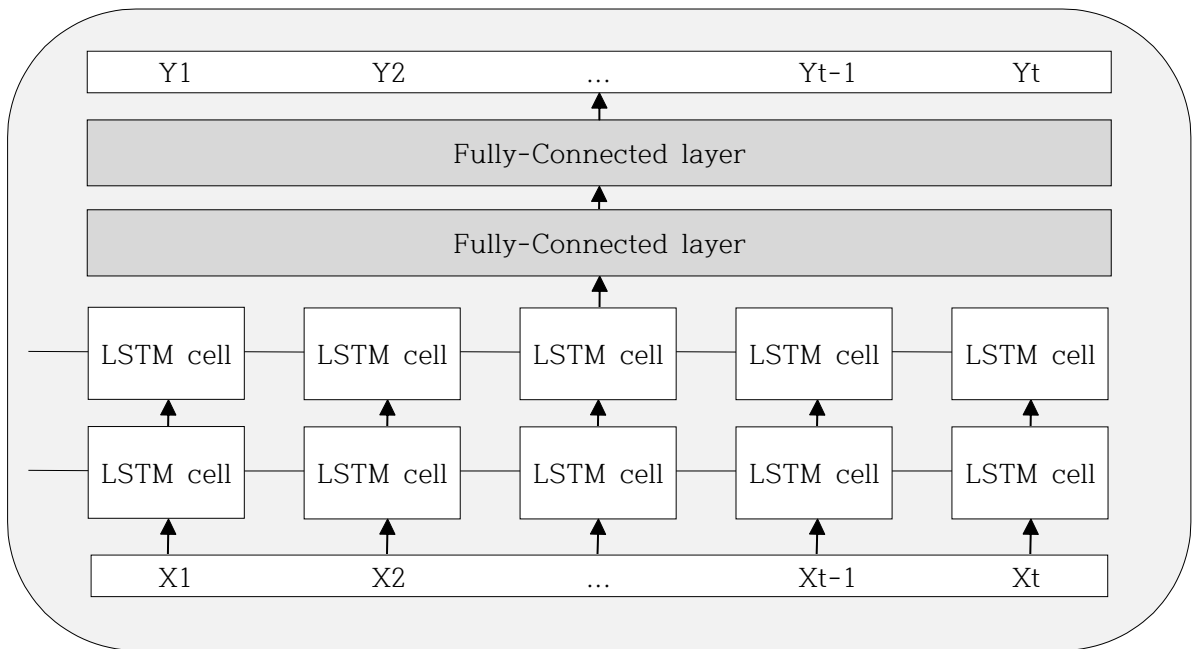
- Input gate: 현재 정보를 기억하기 위한 과정으로, 현재의 cell state 값에 얼마를 더할지를 결정함 (식 1-6)

$$\begin{aligned} i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t &= \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \\ C_t &= f_t * C_{t-1} + i_t * \tilde{C}_t \end{aligned} \quad (\text{식 1-6})$$

- Output gate: 최종적으로 얻어진 cell state 값을 얼마나 출력할지 결정하는 과정임 (식 1-7)

$$\begin{aligned} o_t &= \sigma(W_o[h_{t-1}, x_t] + b_o) \\ h_t &= o_t * \tanh(C_t) \end{aligned} \quad (\text{식 1-7})$$

- KTX 수송수요 예측을 위한 최종 딥러닝 기반 시계열 모형은 <그림 2-8>와 같은 형태로 구성됨
- 완전연결계층(Fully-Connected layer): 딥러닝 모형의 이전 은닉층과 다음 은닉층의 모든 뉴런(neuron)을 연결한 것을 의미하며, 본 연구에서는 KTX 수송수요 예측을 위하여 LSTM 모형과 결합하여 시간적 정보를 통합하는 과정에 사용됨



<그림 2-8> 딥러닝 기반 시계열 예측 모형 개요도

2.4 수요 예측 방법 정립

가. 자료 입력

- 부적절 자료 제거
 - 자료 집계 이전에 자료가 지나치게 적거나 고속열차에 해당하지 않는 자료 등을 사전에 제거하는 과정을 수행함
- KTX 수요 예측 시계열 모형 분석을 위한 입력자료 생성
 - KTX 노선별 월별 수요 예측을 위하여 입력 자료를 월 단위의 노선별 수송인원으로 집계
 - KTX 정차역별 월별 수요 예측을 위하여 입력 자료를 월 단위의 주요 정차역별(즉, 서

- 울역/동대구역/부산역) 수송인원으로 집계
- 집계된 자료와 부합하는 이외 다른 변수(즉, 운행횟수/공급좌석수/승차율/이용율)를 추가하여 최종 입력자료 생성

나. 시계열 모형 적합 및 검증 단계

- 수요 예측을 위한 입력 자료의 집계 단위는 노선별/정차역별 수송인원으로 구성되어 있으며, 이를 이용한 수요예측 모형을 구성함
- KTX 수송수요 시계열 예측모형 적합 및 검증
 - <그림 2-9>의 “시계열 모형 수송수요 예측 과정”에 예측 과정이 도시되어 있음
 - 가. 자료 입력 단계에서 생성된 노선별/정차역별 월별 수송실적 자료를 이용함
 - 2015년 1월부터 2023년 4월 기간의 자료를 이용하여 통계 기반 시계열 모형(즉, ARIMA, Prophet)과 딥러닝 기반 시계열 모형(즉, LSTM)의 적합 및 검증을 시행함
 - 시계열 모형의 적합 기간은 2015년 1월부터 2022년 4월까지이며 검증기간은 2022년 5월부터 2023년 4월까지이며, 해당 기간 동안의 예측오차가 최소가 되는 모형을 선정함

다. 최종 시계열 예측 모형 적합 및 장래 예측

- KTX 수송수요 시계열 예측모형 최종 적합 및 장래 예측
 - <그림 2-9>의 “시계열 모형 수송수요 예측 과정”에 예측 과정이 도시되어 있음
 - 2015년 1월부터 2022년 4월까지 구축된 자료를 이용하여 최종 시계열 모형을 적합함
 - 최종 적합 된 시계열 모형을 이용하여 2024년 01월-12월 기간의 수송수요를 예측함



<그림 2-9> 시계열 모형 수송수요 예측 과정

제 3 장 예측 결과 검증

3.1 KTX 노선별 수송수요 예측 결과 검증

3.2 KTX 주요 정차역별 수송수요 예측 결과 검증

제3장 예측 결과 검증

3.1 KTX 노선별 수송수요 예측 결과 검증

- 시계열 모형을 이용한 노선별 예측치와 실제 집계된 수송실적과 비교한 예측 결과 검증은 다음과 같음
 - 2015년 1월 ~ 2022년 4월 기간 동안의 실적치로 학습을 진행함
 - 2022년 5월 ~ 2023년 4월 기간 동안의 예측치와 실적치를 총 5가지 성능 지표 (R-Squared, RMSE, MSE, MAE, MAPE)을 이용해 비교하였음
 - 노선별 수송수요 예측 결과를 확인하면 통계 기반 시계열 분석 모형인 ARIMA, Prophet에 비해 딥러닝 기반 시계열 모형 LSTM이 R2 score 0.9765로 더 높아 통계 기반 모형에 비해 우수함을 입증함
 - 월별 오차율은 대체로 노선별은 평균 -0.61%로 나옴

구분		노선별					정차역별				
		R2	RMSE	MSE	MAE	MAPE	R2	RMSE	MSE	MAE	MAPE
2022년 5월 ~ 2023년 4월	ARIMA	0.8720	352313	124124 766977	252334	0.1973	-	-	-	-	-
	Prophet	0.9509	219213	480545 42529	172278	0.1608 92	-	-	-	-	-
	LSTM (Ours)	0.9765	150943	227838 19333	81587	0.0686	0.9354	86128	741802 9877	67093	0.1042 28

<표 3-1> 2022년 5월 ~ 2023년 4월 노선별 수송수요 예측결과 검증

구분		노선별 열차 월별 평균 수송 수요				정차역별 열차 월별 평균 수송 수요			
		예측치	실적치	오차	오차율	예측치	실적치	오차	오차율
2022년	5월	910834	1199648	-288814	-24.07	474084	711696	-237612	-33.39
	6월	1084717	1135102	-50385	-4.44	597454	685170	-87716	-12.8
	7월	1162235	1162640	-405	-0.03	665244	711953	-46709	-6.56
	8월	1244053	1146179	97874	8.54	702821	713162	-10341	-1.45
	9월	1250190	1158040	92150	7.96	709261	685449	23812	3.47
	10월	1316167	1315647	520	0.04	695513	809268	-113755	-14.06
	11월	1308977	1216693	92284	7.58	738988	743201	-4213	-0.57
2023년	12월	1299215	1259158	40057	3.18	720927	780371	-59444	-7.62
	1월	1235692	1201802	33890	2.82	701897	734445	-32548	-4.43
	2월	1134641	1174115	-39474	-3.36	686647	729890	-43243	-5.92
	3월	1178525	1205128	-26603	-2.21	687819	718372	-30553	-4.25
	4월	1217766	1259007	-41241	-3.28	684411	762387	-77976	-10.23
평균절대오차/ 평균절대오차율		1195251	1202763	-7512.3	-0.61	672088.8	732113.7	-60024.8	-8.15

<표 3-2> LSTM 모델을 활용해 얻은 월별 수송수요 총량 예측치와 실적치 비교

3.1.1 ARIMA: KTX 수송수요 예측 결과 검증

- 통계 기반 시계열 모형 ARIMA의 KTX 수송수요 예측 결과 비교
 - 전체적인 증감 경향을 따라가지 못하는 모습을 보이며 2023년 2월 예측 시 수송 수요가 크게 감소하는 방향으로 예측함 (<그림 3-1>, <그림 3-2>, <그림 3-3>, <그림 3-4>, <그림 3-5> 참고)



<그림 3-1> ARIMA의 경부선 예측치와 실측치 비교



<그림 3-2> ARIMA의 호남선 예측치와 실측치 비교



<그림 3-3> ARIMA의 경전선 예측치와 실측치 비교



<그림 3-4> ARIMA의 전라선 예측치와 실측치 비교



<그림 3-5> ARIMA의 동해선 예측치와 실측치 비교

3.1.2 Prophet: KTX 수송수요 예측 결과 검증

- 통계 기반 시계열 모형 Prophet의 KTX 수송수요 예측 결과 비교
 - ARIMA와 비교하였을 때 상대적으로 경향성을 따라 가는 모습을 보이나 5개의 주운행선 모두 예측 경향이 비슷함을 확인할 수 있음 (<그림 3-6>, <그림 3-7>, <그림 3-8>, <그림 3-9>, <그림 3-10> 참고)



<그림 3-6> Prophet의 경부선 예측치와 실측치 비교



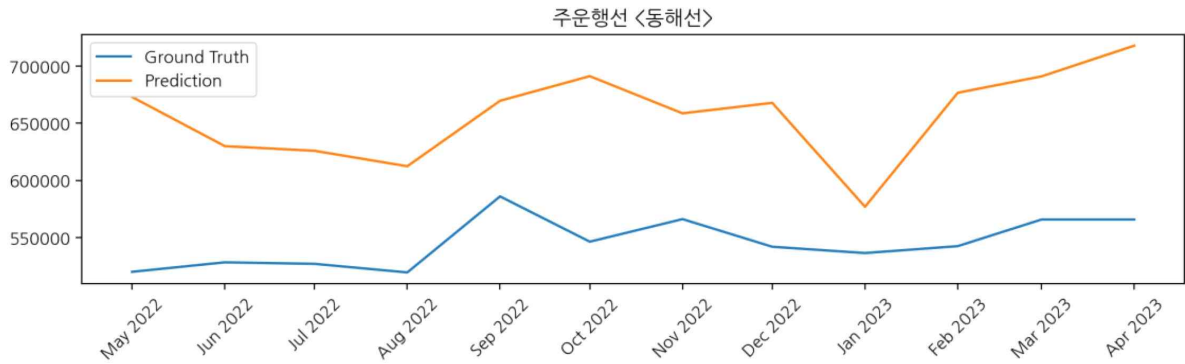
<그림 3-7> Prophet의 호남선 예측치와 실측치 비교



<그림 3-8> Prophet의 경전선 예측치와 실측치 비교



<그림 3-9> Prophet의 전라선 예측치와 실측치 비교



<그림 3-10> Prophet의 동해선 예측치와 실측치 비교

3.1.4 LSTM: KTX 수송수요 예측 결과 검증

- 딥러닝 기반 시계열 모형 LSTM의 KTX 수송수요 예측 결과 비교
 - 이전의 통계 기반 시계열 모형인 ARIMA 및 Prophet과 비교했을 때 딥러닝 기반 시계열 모형은 5개의 주운행선 간의 특징을 상대적으로 우수하게 잡아내며 예측값이 다른 모델에 비해 실제 수송수요와 비슷한 경향을 따라감 (<그림 3-11>, <그림 3-12>, <그림 3-13>, <그림 3-14>, <그림 3-15> 참고)



<그림 3-11> LSTM의 경부선 예측치와 실측치 비교



<그림 3-12> LSTM의 호남선 예측치와 실측치 비교



<그림 3-13> LSTM의 경전선 예측치와 실측치 비교



<그림 3-14> LSTM의 전라선 예측치와 실측치 비교



<그림 3-15> LSTM의 동해선 예측치와 실측치 비교

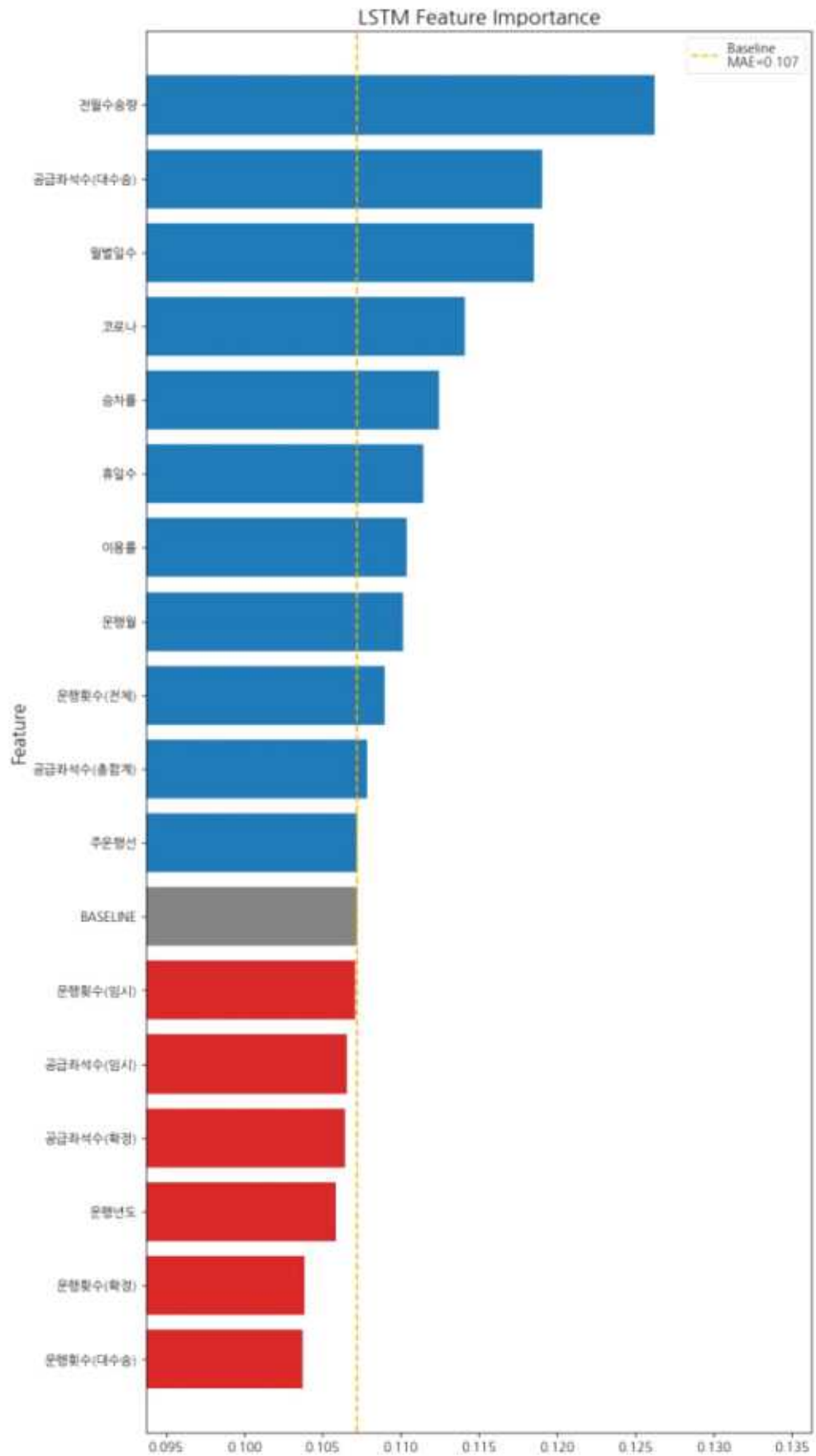
- 또한 딥러닝 기반 시계열 모형 LSTM이 KTX 수송수요 예측에 있어서 고려한 변수의 영향도는 <표 3-4>와 같이 도출됨 (<그림 3-16> 참고)

변수	Feature Importance
전월수송량	0.126245
공급좌석수(대수송)	0.119025
월별일수	0.118502
코로나	0.114125
승차율	0.112429
휴일수	0.11147
이용률	0.110395
운행월	0.110182
운행횟수(전체)	0.108994
공급좌석수(총합계)	0.107855
주운행선	0.107215
BASELINE	0.107215
운행횟수(임시)	0.107105
공급좌석수(임시)	0.106562
공급좌석수(확정)	0.106425
운행년도	0.105833
운행횟수(확정)	0.103871
운행횟수(대수송)	0.103744

<표 3-3> LSTM 변수 영향도

- LSTM 변수 영향도를 측정 결과 전월수송량(0.126245), 공급좌석수(대수송)(0.119025), 월별일수(0.118502), 코로나(0.114125), 승차율(0.112429) 순으로 KTX 수송수요 예측 문제에 있어 가장 영향력 있는 변수 5개로서 고려되었음

- 종합적으로 전월수송량(0.126245)이 KTX 수송수요 예측에 있어 가장 영향력 있는 변수임을 확인함
- 그 외 영향력 있는 변수로는 이용률, 운행횟수가 있었으며 운행횟수(확정)(0.103871)과 공급좌석수(임시)(0.106562)는 상대적으로 영향력이 덜함을 확인함



<그림 3-16> LSTM 변수 영향도

3.1.4 승차율 및 이용률 정보 효과 탐색

- 승차율, 이용률은 KTX 수송수요 예측에 있어 중요한 영향을 미치는 것을 <그림 3-16>을 통해 알 수 있음. 따라서 해당 변수를 구성 방식에 따른 모델 검증 값 변화를 알아보았음
 - 동일한 해의 당월 승차율 및 이용률을 사용하여 최종 수송수요를 예측하는 방법과, 그 전 달의 승차율 및 이용률을 사용하여 수송수요를 예측하는 방법 두 가지 측면에서 살펴봄
- 승차율, 이용률은 KTX 수송수요 예측에 있어 중요한 영향을 미치는 것을 <그림 3-16>을 통해 알 수 있음. 따라서 해당 변수를 구성 방식에 따른 모델 검증 값 변화를 알아보았음
 - 통계 기반 시계열 모형인 ARIMA는 기존 0.8727 R2 score에서 0.8694로 하락하였으며, Prophet은 기존 0.9509 R2 score에서 0.8423로 가장 크게 하락함
 - 딥러닝 기반 시계열 모형인 LSTM은 기존 0.9765 R2 score에서 0.9721로 하락함
 - 종합적으로 승차율 및 이용률 변수를 해당월을 사용하는 경우보다 전월 승차율 및 이용률을 사용하는 경우 모든 시계열 예측 모형에서 성능이 하락함을 확인할 수 있음

구분		노선별				
		R2	RMSE	MSE	MAE	MAPE
해당 년도 승차율, 이용률 사용 시(원안)	ARIMA	0.8727	351373	124124766977	249129	0.190914
	Prophet	0.9509	219213	48054542529	172278	0.160892
	LSTM(Ours)	0.9765	150943	22783819333	81587	0.0686
이전 달의 승차율, 이용률 사용 시	ARIMA	0.8694	355842	126624125073	262480	0.2106
	Prophet	0.8423	393017	154462421864	260202	0.1882
	LSTM(Ours)	0.9721	164559	27079955319	98937	0.0775

<표 3-4> 당월 이용률 및 승차율과 전월 이용률 및 승차율을 사용했을 때 검증 결과 변화

3.2 KTX 주요 정차역별 수송수요 예측 결과 검증

- 딥러닝 기반 시계열 모형 LSTM을 이용한 정차역별 예측치와 실제 집계된 수송실적과 비교한 예측 결과 검증은 <표 3-5>와 같음
 - 2018년 1월 ~ 2022년 4월 기간 동안의 실적치를 활용하여 딥러닝 기반 시계열 모형 LSTM의 적합을 진행함

- 2022년 5월 ~ 2023년 4월 기간 동안의 예측치와 실적치를 총 5가지 성능 지표 (R-Square, RMSE, MSE, MAE, MAPE)을 이용해 비교하였음
- <표 3-6>을 통하여 전체적으로 역별 오차율은 평균 -8% 로 나옴을 확인함

구분		노선별					정차역별				
		R2	RMSE	MSE	MAE	MAPE	R2	RMSE	MSE	MAE	MAPE
2022년 5월 ~ 2023년 4월	ARIMA	0.8727	351373	124124 766977	249129	0.19091	-	-	-	-	
	Prophet	0.9509	219213	480545 42529	172278	0.16089	-	-	-	-	
	LSTM (Ours)	0.9765	150943	227838 19333	81587	0.0686	0.93535	86128.0	741802 9877	67093.1	0.10423

<표 3-5> 2022년 5월 ~ 2023년 4월 노선별 수송수요 예측결과 검증

구분		노선별 열차 월별 평균 수송 수요				정차역별 열차 월별 평균 수송 수요			
		예측치	실적치	오차	오차율	예측치	실적치	오차	오차율
2022년	5월	910834	1199648	-288814	-24.07	474084	711696	-237612	-33.39
	6월	1084717	1135102	-50385	-4.44	597454	685170	-87716	-12.8
	7월	1162235	1162640	-405	-0.03	665244	711953	-46709	-6.56
	8월	1244053	1146179	97874	8.54	702821	713162	-10341	-1.45
	9월	1250190	1158040	92150	7.96	709261	685449	23812	3.47
	10월	1316167	1315647	520	0.04	695513	809268	-113755	-14.06
	11월	1308977	1216693	92284	7.58	738988	743201	-4213	-0.57
2023년	12월	1299215	1259158	40057	3.18	720927	780371	-59444	-7.62
	1월	1235692	1201802	33890	2.82	701897	734445	-32548	-4.43
	2월	1134641	1174115	-39474	-3.36	686647	729890	-43243	-5.92
	3월	1178525	1205128	-26603	-2.21	687819	718372	-30553	-4.25
	4월	1217766	1259007	-41241	-3.28	684411	762387	-77976	-10.23
평균절대오차/ 평균절대오차율		1195251	1202763	-7512	-0.60	672088.8	732113.7	-60024.8	-8.15

<표 3-6> LSTM 모델을 활용해 얻은 월별 수송수요 총량 예측치와 실적치 비교

제 4 장 수송수요 예측 결과

4.1 2024년도 KTX 노선별 수송수요 예측

4.2 2024년도 KTX 주요 정차역별 수송수요 예측

제4장 수송수요 예측 결과

4.1 2024년도 KTX 노선별 수송수요 예측 결과

- 본 연구에서 제안하는 딥러닝 기반 시계열 모델을 활용하여 예측한 2024년도 KTX 노선별 월별 수송인원 예측값은 <표 4-1>, <표 4-2>, <표 4-3>, <표 4-4>, <표 4-5>와 같음
 - 2024년 경부선에 대한 평균 수송인원은 3323473으로 2022년 평균 수송인원의 2,762,553에 비해 17.4% 증가함
 - 2024년 경전선에 대한 평균 수송인원은 700456으로 2022년 평균 수송인원의 581,832에 비해 17.7% 증가함
 - 2024년 동해선에 대한 평균 수송인원은 579807으로 2022년 평균 수송인원의 481,971에 비해 17.1% 증가함
 - 2024년 전라선에 대한 평균 수송인원은 763824로 2022년 평균 수송인원의 626,773에 비해 18.72% 증가함
 - 2024년 호남선에 대한 평균 수송인원은 1086072으로 2022년 평균 수송인원의 865,356에 비해 21.34% 증가함

구분			노선별		
주운 행선	예측해	월	2022년	2023~2024년	증가율
경부선	2023	5	3104963	3147931	1.38
		6	2952737	3211557	8.77
		7	3041059	3256460	7.08
		8	3004164	3286810	9.41
		9	2993883	3258361	8.83
		10	3453559	3512441	1.7
		11	3194372	3561796	11.5
	2024	12	3326734	3448755	3.67
		1	2078548	3260394	56.86
		2	1753695	2843824	62.16
		3	1819050	3104456	70.66
		4	2427868	3279061	35.06
		5	3104963	3392522	9.26
		6	2952737	3391473	14.86
		7	3041059	3362209	10.56
		8	3004164	3218436	7.13
		9	2993883	3289960	9.89
		10	3453559	3535814	2.38
		11	3194372	3615734	13.19
		12	3326734	3491473	4.95
		평균	2911105	3323473	17.46

<표 4-1> 최종 모델의 경부선 2024 수송 수요 예측 결과

구분			노선별		
주요행선	예측해	월	2022년	2023~2024년	증가율
경전선	2023	5	655310	698784	6.63
		6	612109	677061	10.61
		7	625831	669508	6.98
		8	606413	685675	13.07
		9	640597	690679	7.82
		10	712557	716206	0.51
		11	658463	720963	9.49
	2024	12	680136	710390	4.45
		1	460220	676737	47.05
		2	381172	623921	63.68
		3	413835	680524	64.44
		4	535344	728810	36.14
		5	655310	738395	12.68
		6	612109	720221	17.66
		7	625831	697111	11.39
		8	606413	679907	12.12
		9	640597	706590	10.3
		10	712557	725743	1.85
		11	658463	737471	12
		12	680136	724433	6.51
		평균	608670	700456	17.77

<표 4-2> 최종 모델의 경전선 2024 수송 수요 예측 결과

구분			노선별		
주운행선	예측년도	월	2022년	2023~2024년	증가율
동해선	2023	5	542201	572970	5.67
		6	520044	576075	10.77
		7	528363	571867	8.23
		8	527032	576323	9.35
		9	519510	577760	11.21
		10	586066	594098	1.37
		11	546410	610012	11.64
	2024	12	566167	604045	6.69
		1	369562	560994	51.8
		2	318196	461869	45.15
		3	332836	549460	65.08
		4	427268	583173	36.49
		5	542201	613195	13.09
		6	520044	600194	15.41
		7	528363	569111	7.71
		8	527032	569902	8.13
		9	519510	580812	11.8
		10	586066	594803	1.49
		11	546410	624404	14.27
		12	566167	605091	6.88
		평균	505972	579807	17.11

<표 4-3> 최종 모델의 동해선 2024 수송 수요 예측 결과

구분			노선별		
주운행선	예측년도	월	2022년	2023~2024년	증가율
전라선	2023	5	707716	762482	7.74
		6	665053	746291	12.22
		7	693723	756479	9.05
		8	686650	769583	12.08
		9	671762	759542	13.07
		10	753752	810314	7.5
		11	692262	780919	12.81
	2024	12	717034	783471	9.27
		1	485751	718607	47.94
		2	422889	606049	43.31
		3	445316	760362	70.75
		4	579374	769868	32.88
		5	707716	809126	14.33
		6	665053	771878	16.06
		7	693723	743229	7.14
		8	686650	762015	10.98
		9	671762	766238	14.06
		10	753752	811130	7.61
		11	692262	801602	15.79
		12	717034	787310	9.8
		평균	655461	763824	18.72

<표 4-4> 최종 모델의 전라선 2024 수송 수요 예측 결과

구분			노선별		
주운행선	예측해	월	2022년	2023~2024년	증가율
호남선	2023	5	988053	1058911	7.17
		6	925571	1079342	16.61
		7	924227	1086830	17.59
		8	906636	1117294	23.24
		9	964448	1086457	12.65
		10	1072303	1146303	6.9
		11	991958	1157076	16.65
	2024	12	1005722	1123958	11.76
		1	662038	1025065	54.83
		2	553406	822054	48.54
		3	605235	819204	35.35
		4	784679	1010144	28.73
		5	988053	1144847	15.87
		6	925571	1174501	26.89
		7	924227	1152995	24.75
		8	906636	1072807	18.33
		9	964448	1121725	16.31
		10	1072303	1168740	8.99
		11	991958	1193747	20.34
		12	1005722	1159448	15.29
		평균	908159	1086072	21.34

<표 4-5> 최종 모델의 호남선 2024 수송 수요 예측 결과



<그림 4-2> 최종 모델의 경부선 2024 수송 수요 예측 결과



<그림 4-1> 최종 모델의 경전선 2024 수송 수요 예측 결과



<그림 4-3> 최종 모델의 동해선 2024 수송 수요 예측 결과



<그림 4-4> 최종 모델의 전라선 2024 수송 수요 예측 결과



<그림 4-5> 최종 모델의 호남선 2024 수송 수요 예측 결과

4.2 2024년도 KTX 주요 정차역별 수송수요 예측 결과

- 본 연구에서 제안하는 딥러닝 기반 시계열 모델을 활용하여 예측한 2024년도 KTX 정차역별 수송인원 예측값은 <표 4-6>, <표 4-7>, <표4-8>와 같음
 - 2024년 서울역에 대한 평균 월별 고속열차 수송인원으로 1248616 명을 예측함
 - 2024년 서울역에 대한 평균 월별 고속열차 수송인원은 가장 적게는 9월에 1203920명을, 가장 많게는 11월에 1374655명이 예상됨
 - 2024년 부산역에 대한 평균 월별 고속열차 수송인원으로 485352명의 수송인원을 예측함
 - 2024년 부산역에 대한 평균 월별 고속열차 수송인원은 가장 적게는 10월에 471198 명이, 가장 많게는 11월에 505195명이 예상됨
 - 2024년 동대구역에 대한 평균 월별 고속열차 수송인원으로 441730명을 예측함

- 2024년 동대구역에 대한 평균 월별 고속열차 수송인원은 가장 적게는 1월에 427136 명이 가장 많게는 11월에 467937명이 예상됨

출발역	연도	월	2024예측
서울	2024	1	1210291
	2024	2	1236235
	2024	3	1236636
	2024	4	1248424
	2024	5	1274289
	2024	6	1259174
	2024	7	1221929
	2024	8	1257167
	2024	9	1203921
	2024	10	1228535
	2024	11	1374655
	2024	12	1232138
평균			1248616

<표 4-6> 최종 모델의 서울역 2024년 고속열차 수송 수요 예측 결과

출발역	연도	월	2024예측
부산	2024	1	486879
	2024	2	476937
	2024	3	492874
	2024	4	480089
	2024	5	484555
	2024	6	485210
	2024	7	478337
	2024	8	483735
	2024	9	491042
	2024	10	471198
	2024	11	505195
	2024	12	488178
평균			485352

<표 4-7> 최종 모델의 부산역 2024년 고속열차 수송 수요 예측 결과

출발역	연도	월	2024예측
동대구	2024	1	427136
	2024	2	435857
	2024	3	428857
	2024	4	421963
	2024	5	431972
	2024	6	454796
	2024	7	436179
	2024	8	450584
	2024	9	448567
	2024	10	464393
	2024	11	467937
	2024	12	432529
평균			441730

<표 4-8> 최종 모델의 동대구역 2024년 고속열차 수송 수요 예측 결과

제 5 장 결론

5.1 KTX 수송수요 예측 결론

5.2 향후 연구

제5장 결 론

5.1 KTX 수송수요 예측 결론

- 2024년 KTX 노선별 수송수요 예측 분석 결과 전체적으로 KTX 수송수요가 2022년 대비 월별 비슷하거나 증가하는 추세로 변화할 것으로 예측되었음
 - 2024년 월별 수송인원은 2022년 동월대비 경부선 평균 17.46%, 경전선 평균 17.77%, 동해선 평균 17.11%, 전라선 평균 수송인원은 최대 18.72%, 호남선 평균 21.34% 증가하는 것으로 예상됨
- 본 연구에서 제안하는 딥러닝 기반 시계열 모델을 활용하여 예측한 2024년도 KTX 정차역별 수송인원은 다음과 같이 예측되었음
 - 2024년 서울역에 대한 평균 월별 수송인원은 가장 적게는 9월에 1203920 인원이, 가장 많게는 11월에 1374655 인원이 예상되며, 부산역에 대한 평균 월별 수송인원은 가장 적게는 10월에 471198 인원이, 가장 많게는 11월에 505195 인원이 예상되며, 마지막으로 동대구역에 대한 평균 월별 수송인원은 가장 적게는 1월에 427136 인원이 가장 많게는 11월에 467937 인원이 예상됨

5.2 향후 연구

- KTX 노선별/정차역별 딥러닝 기반 시계열 예측 모형의 고도화 작업이 필요함
 - 노선별 딥러닝 기반 시계열 모형에 새롭게 반영되는 개입 효과를(예: 날씨 등과 같은 외생변인으로부터 받는 영향 등) 설명할 수 있는 지속적인 모형 개선 작업이 필요함

【참고문헌】

- 김범승, “계절 ARIMA 모형을 이용한 여객수송수요 예측: 중앙선을 중심으로”, 한국철도학회논문집, 제17권 제4호, pp.307-312, 2014.
- Shumway, R. H., Stoffer, D. S., Shumway, R. H., & Stoffer, D. S. (2017). ARIMA models. Time series analysis and its applications: with R examples. 75-163.
- Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018, December). A comparison of ARIMA and LSTM in forecasting time series. In 2018 17th IEEE international conference on machine learning and applications (ICMLA) (pp. 1394-1401). IEEE.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time Series Analysis: Forecasting and Control. John Wiley & Sons.
- Toharudin, T., Pontoh, R. S., Caraka, R. E., Zahroh, S., Lee, Y., & Chen, R. C. (2023). Employing long short-term memory and Facebook prophet model in air temperature forecasting. Communications in Statistics-Simulation and Computation, 52(2), 279-290.
- Prophet, Open source software released by Facebook’s Core Data Science team, <https://facebook.github.io/prophet/>
- Jha, B. K., & Pande, S. (2021, April). Time series forecasting model for supermarket sales using FB-prophet. In 2021 5th International Conference on Computing Methodologies and Communication (ICCMC) (pp. 547-554). IEEE.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.