

# Temporal Fusion Transformers and Deep Learning Methods for Multi-Horizon Time Series Forecasting

InKyung Kim<sup>†</sup> · DaeHee Kim<sup>††</sup> · Jaekoo Lee<sup>†††</sup>

## ABSTRACT

Given that time series are used in various fields, such as finance, IoT, and manufacturing, data analytical methods for accurate time-series forecasting can serve to increase operational efficiency. Among time-series analysis methods, multi-horizon forecasting provides a better understanding of data because it can extract meaningful statistics and other characteristics of the entire time-series. Furthermore, time-series data with exogenous information can be accurately predicted by using multi-horizon forecasting methods. However, traditional deep learning-based models for time-series do not account for the heterogeneity of inputs. We proposed an improved time-series predicting method, called the temporal fusion transformer method, which combines multi-horizon forecasting with interpretable insights into temporal dynamics. Various real-world data such as stock prices, fine dust concentrates and electricity consumption were considered in experiments. Experimental results showed that our temporal fusion transformer method has better time-series forecasting performance than existing models.

Keywords : Time Series, Multi-variate Data Analysis, Multi-horizon Forecasting, Deep Learning, Neural Networks

# Temporal Fusion Transformers와 심층 학습 방법을 사용한 다층 수평 시계열 데이터 분석

김 인 경<sup>†</sup> · 김 대 희<sup>††</sup> · 이 재 구<sup>†††</sup>

## 요 약

시계열 데이터는 주식, IoT, 공장 자동화와 같은 다양한 실생활에서 수집되고 활용되고 있으며, 정확한 시계열 예측은 해당 분야에서 운영 효율성을 높일 수 있어서 전통적으로 중요한 연구 주제이다. 전반적인 시계열 데이터의 향상된 특성을 추출할 수 있는 대표적인 시계열 데이터 분석 방법인 다층 수평 예측은 최근 부가적 정보를 포함하는 시계열 데이터에 내재한 이질성(heterogeneity)까지 포괄적으로 분석에 활용하여 향상된 시계열 예측한다. 하지만 대부분의 심층 학습 기반 시계열 분석 모델들은 시계열 데이터의 이질성을 반영하지 못했다. 따라서 우리는 잘 알려진 temporal fusion transformers 방법을 사용하여 실생활과 밀접한 실제 데이터를 이질성을 고려한 다층 수평 예측에 적용하였다. 결과적으로 주식, 미세먼지, 전기 소비량과 같은 실생활 시계열 데이터에 적용한 방법이 기존 예측 모델보다 향상된 정확도를 가짐을 확인할 수 있었다.

키워드 : 시계열 분석, 다변량 데이터 분석, 다층 수평 예측, 심층 학습, 깊은 인공신경망

## 1. 서 론

시계열 예측 문제는 과거의 관측값을 사용하여 미래를 알

거나 정확하게 예측할 수 있는지에 대해서 다룬다. 경제, 의료 등 다양한 분야의 많은 데이터가 시계열 형태를 가지고 있는 만큼 정확한 예측은 운영 효율성을 높이는 데 도움을 준다. 다층 수평(multi-horizon) 예측은 기계 학습에서 시계열 데이터를 사용한 중요한 문제 중 하나이다. 다층 수평 예측은 전반적인 경향에 대한 짐작을 사용자들에게 제공하고 이를 통해서 사용자들은 미래에 대한 계획을 세우는 데에 도움을 얻을 수 있다.

하지만 대부분의 시계열 데이터는 다차원의 형태를 가지며 다양한 정보를 포함하기 때문에 전통적인 방법을 사용하면 정확한 예측이 어렵다. 이는 포함된 데이터의 이질성으로 인해 상호 작용에 대한 정보가 부족하기 때문이다. 예를 들어,

※ 이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(2020-0-01826, AI 기반 선도적 실전문제 해결 연구인재 양성).

※ 이 논문은 2021년 한국정보처리학회 춘계학술발표대회에서 "Temporal Fusion Transformer 모델을 활용한 다층 수평 시계열 데이터 분석"의 제목으로 발표된 논문을 확장한 것임.

† 준 회 원 : 국민대학교 컴퓨터공학과 석사과정

†† 비 회 원 : 국민대학교 컴퓨터공학과 석사과정

††† 비 회 원 : 국민대학교 SW학부 조교수

Manuscript Received : July 9, 2021

First Revision : August 11, 2021

Accepted : August 26, 2021

\* Corresponding Author : Jaekoo Lee(jaekoo@kookmin.ac.kr)

특정 기업에서 소비자들의 구매 행태를 예측하려고 할 때, 기업은 각 상품에 대한 구매 정보와 함께 구매처, 구매 물품, 구매자에 대한 정보 등 다양한 정보가 필요하다. 하지만 구매 정보는 실수형의 형태로 시간에 따라 데이터가 축적되지만 구매 물품에 대한 정보는 범주형의 형태를 가지며 시간에 따라 축적되는 값이 아니다. 또한 소비자들의 구매에 영향을 주는 공휴일의 경우 날짜의 형태를 가진다. 이처럼 시계열 정보들은 실수형, 범주형 등의 이질적인 형태를 가지며 서로의 관계에 대한 정보를 매우 적게 제공한다. 이러한 이질성이 높은 시계열 데이터의 특성으로 인해 다층 수평 시계열 예측은 매우 어려운 과업 중 하나이다. 최근 깊은 신경망의 발전으로 새로운 방법들이 제안되었지만, 여전히 다양한 형태의 입력값을 활용한 예측에는 한계가 있다.

이러한 한계점을 보완하고자 본 논문에서는 Temporal Fusion Transformers (TFT)[1]와 기존 예측 모델인 Long Short-Term Memory (LSTM)[2], 1 Dimensional (1D) Convolutional Neural Network (CNN)[3]을 실생활과 밀접한 관련이 있는 실제 데이터에 적용하여 예측 성능을 비교하였다<sup>1)</sup>. 실험 결과, TFT 방법이 LSTM이나 1D CNN 방법보다 더 정확한 예측이 가능한 것을 확인하였다.

## 2. 관련 연구

시계열 분석은 다변량(multivariate) 분석과 단변량(univariate) 분석으로 나뉜다. 다변량 분석은 독립 변수로 사용하는 추가적인 시계열 데이터를 고려한 것으로 단변량 분석을 확장한 것이다. 반면 단변량 분석은 과거의 값을 통해서 미래의 관측값을 예측한다. 수평(horizon) 예측은 일반적으로 한 단계 앞선 예측에 중점을 둔다. 반면 다층 수평 예측은 관심 변수를 여러 미래 시간 단계에서 예측한다.

깊은 신경망 이전의 시계열 예측 문제는 통계적 모델을 사용하여 단변량 분석을 진행하였다. 기존의 통계적 모델로는 autoregressive integrated moving average(ARIMA), simple moving average(SMA), exponential smoothing(ETS) 등이 있다. 하지만 위의 모델들은 상대적으로 단순하기 때문에 정확한 단변량 분석에는 한계가 있다[4]. 최근 다층 수평 예측에는 깊은 신경망이 사용되고 있으며 그중 많은 모델이 순환 신경망 구조에 기반한다[5-8]. 특히 최근에는 어텐션(attention) 기반의 방법을 사용한 구조들이 제안되었다[7,8]. 하지만 아직 많은 모델이 데이터의 다양한 형태인 이질성을 고려하지 못하는 한계가 있다[5,6,8]. DSTP-RNN[7]의 경우 여러 종류의 입력값을 고려하지만, 실숫값을 가진 데이터 집합만을 사용하여 실험을 진행하였다.

하지만 시계열 데이터는 실숫값 이외에도 다양한 형태를 띠기 때문에 이를 모두 반영해야 정확한 예측이 가능하다.

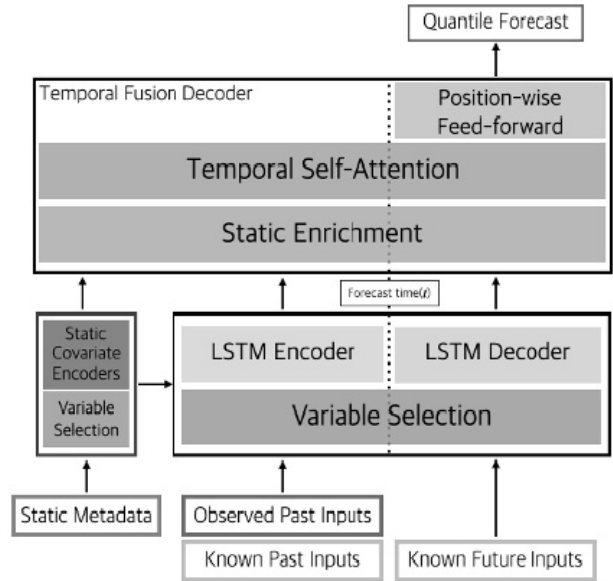


Fig. 1. TFT Method Architecture

TFT 방법은 여러 형태의 값들을 입력값과 다양한 종류의 데이터를 사용할 수 있다. 본 연구에서는 이전 연구들이 데이터의 이질성을 반영하지 못하는 한계점을 해결하기 위해서 TFT 방법을 사용하였다.

## 3. TFT 방법을 활용한 다층 수평 예측

시계열 데이터는 정적 공변량 값, 알려진 미릿값과 외인성의 데이터를 포함하는 복잡한 형태로 이루어져 있으며 이 정보들은 상호 작용하여 형성된다. 예를 들어, 소매업 데이터 [9]는 상품과 가계에 대한 메타데이터(metadata)를 포함하면서 일일 수준의 외인성 시간 변화에 따른 값을 포함한다. 또한 과거와 미래 모두에 알려진 값으로 볼 수 있는 공휴일 정보, 세일 정보 등이 포함된다. 때문에 다층 수평 예측은 다양한 형태의 데이터를 반영할 수 있어야 한다. TFT 방법[1]은 이전 모델들이 가진 한계점을 보완하여 다양한 유형의 입력값을 고려하도록 설계되었다.

### 3.1 Temporal Fusion Transformers 방법

TFT 방법[1]은 어텐션 방법에 기반한 구조이며 다양한 규모의 시간적 관계성을 파악하기 위해서 지역적 처리를 위한 순환 신경망 층과 장기 의존성을 위한 셀프-어텐션(self-attention) 층을 포함한다. 또한 유의미한 특징 추출을 위한 변수 선택(variable selection) 층과 중요하지 않은 특징을 억제하기 위한 개폐(gating) 층을 사용한다. 그래서 다양한 시나리오에서 좋은 성능을 보이며 분위수 손실 함수를 사용하여 결과의 q-분위수를 예측한다. 전반적인 구조는 Fig. 1을 통해서 확인할 수 있다.

TFT 방법은 다음과 같이 모델을 구성한다. 먼저 주어진

1) 관련된 코드는 <https://github.com/IKKIM00/multi-horizon-forecasting-comparison-between-TFT-and-DL-methods.git> 에서 확인할 수 있습니다.

시계열 데이터의 고유한 존재  $i$ 가 존재한다고 가정한다. 예를 들면  $i$ 는 소매업에서는 개별 가게를 의미한다.

$$\hat{y}_i(q, t, \tau) = f(q, \tau, y_{i, t-k:t}, z_{i, t-k:t}, x_{i, t-k:t+\tau}, s_i) \quad (1)$$

각 존재  $i$ 는 정적 공변량 값  $s_i$ 와 입력값  $x_{i,t}$ , 정수값  $y_{i,t}$  모두와 연관이 있다. 시간에 의존적인 입력값  $x_{i,t} = [z_{i,t}^T, x_{i,t}^T]$ 은 매시간 측정되고 미리 알 수 없는 관측된 입력값  $z_{i,t}$ 와 사전에 결정되는 알려진 입력값  $x_{i,t}$ 로 나누어진다.  $\tau$ 는 분위수를 의미한다. 모델의 입력과 각각의 분위수 예측은 위의 Equation (1)과 같이 이루어진다.

모델에 입력값이 들어오면 가장 먼저 변수 선택 층을 통과하고 이 과정을 통해서 유의미한 정보를 추출한다. 추출한 정보는 LSTM 인코더(encoder)와 LSTM 디코더(decoder)로 이루어진 시퀀스-투-시퀀스(sequence-to-sequence) 층의 입력값으로 사용한다. 시퀀스-투-시퀀스 층의 초기 상태 값으로는 정적 공변량 값  $s_i$ 의 컨텍스트 벡터(context vector)를 사용하여 관측된 입력값  $z_{i,t}$ 와 알려진 입력값  $x_{i,t}$ 과의 시간적 특징(temporal feature)를 생성한다. 이후 정적 강화(static enrichment) 층에서 정적 메타데이터의 시간적 특징을 강화하고 시간적 셀프 어텐션(temporal self-attention) 층을 통해서 인과적 정보의 흐름과 장기 의존성을 보존한다. 마지막으로 위치별 피드포워드(position-wise feed-forward) 층을 통해서 추가적인 비선형성을 적용한다.

### 3.2 데이터 설명

실험은 실생활과 밀접한 관련이 있는 세 개의 공개 데이터인 Carriage Services 기업[10]의 주식 데이터, 베이징의 미세먼지 데이터[11], 전기 변압기 데이터[12]를 사용하여 실험을 진행하였다. 세 데이터 모두 갑작스럽게 변화할 수 있는 단기적인 속성과 장기적인 속성을 모두 파악해야 정확한 예측이 가능하다.

주식 데이터는 Yahoo finance에서 제공한 Carriage Services 기업[10]의 데이터를 사용하였다. Yahoo finance는 기업별로 하루 단위의 시가, 고가, 저가, 종가, 거래량, 수정 증가에 대한 값을 제공한다. 실험에는 1996년 8월부터 2021년 4월까지의 시가, 고가, 저가, 종가와 거래량을 사용하여 진행하였다.

베이징 미세먼지 데이터[11]는 데이터의 측정 시점, 이슬점, 온도, 기압, 풍향, 풍속, 적설 시간, 누적 강우 시간, 미세먼지 농도에 대한 정보를 포함한다. 데이터는 2010년 1월부터 2014년 12월까지 한 장소에서 시간 단위로 측정되었다. 실험에는 위의 모든 데이터를 사용하여 미세먼지 농도를 예측하도록 설정하였다.

전기 변압기 데이터[12]는 2016년 7월부터 2018년 7월까지 변압기를 사용하여 측정되었으며 변압기 기름 온도와 외부 전기 전원 부하에 관한 6가지 데이터를 포함한다. 데이터

는 두 장소에서 15분 단위로 측정되었다. 실험에는 주어진 데이터를 모두 사용하여 변압기 기름 온도를 예측하였다.

### 3.3 전처리 방법

미세먼지 데이터[11]의 경우 다음과 같이 전처리하였다. 훈련 집합, 검증 집합, 실험 집합을 각각 2010년 1월 1일~2012년 12월 31일, 2013년 1월 1일~2013년 12월 12일, 2014년 1월 1일~2014년 12월 31일로 설정했다. 데이터가 시간 단위로 수집되었던 점을 활용하여 24시간의 온도, 이슬점, 기압, 풍속, 풍향, 적설 시간, 누적 강우 시간을 사용해서 다음 한 시간의 미세먼지 농도를 예측할 수 있도록 설정했다. TFT 방법에 적용하기 위한 입력값 설정은 다음과 같다. 데이터가 시간 값과 측정된 날짜는 알려진 입력값으로 사용하였다. 정적 공변량 값으로는 데이터가 측정된 장소에 대한 정보를 사용했다. 관측된 입력값으로는 적설 시간, 이슬점, 온도, 풍향, 기압, 풍속, 누적 강우 시간을 모두 사용하였다.

주식 데이터[10]의 경우 훈련 집합, 검증 집합, 실험 집합을 각각 1996년 8월 9일~2011년 12월 31일, 2012년 1월 1일~2016년 12월 31일, 2017년 1월 1일~2021년 4월 7일로 설정했다. 예측 주기는 50일의 데이터를 사용해서 다음 날 증가를 예측하도록 설정하였다. TFT 방법에 적용하기 위한 입력값은 다음과 같이 설정했다. 정적 공변량 값은 하나의 기업 데이터만 사용하였기 때문에 하나의 정수값으로 설정했다. 관측된 입력값으로는 시가, 고가, 저가, 종가, 거래량 값을 사용하였다.

전기 변압기 데이터[12]는 훈련 집합, 검증 집합, 실험 집합으로 2016년 7월 1일~2017년 6월 30일, 2017년 7월 1일~2018년 1월 31일, 2018년 2월~2018년 6월 26일까지의 데이터를 사용하였다. 예측 주기는 5일의 데이터를 사용해서 다음 측정 시기의 기름 온도를 예측하도록 설정하였다. TFT 방법에 적용하기 위해서 정적 공변량 값은 각 변압기의 측정 장소로 설정하였고 관측된 입력값으로는 6가지 유형의 외부 전력 부하 기능 값을 사용하였다.

## 4. 실험 결과

TFT 방법과 함께 기존의 예측 모델로 단층 LSTM 모델 [2], 다층 LSTM 모델 [2], 1D CNN 모델 [3]을 사용했다. 각 모델은 이전 연구에서 제안한 모델 구조를 참고하여 변형하였다. 단층 LSTM 모델과 다층 LSTM 모델의 경우 LSTM 층을 통과한 출력값이 선형 층과 배치 정규화 층을 통과한 후 예측값을 출력하도록 설정하였다. 1D CNN 모델의 경우, 합성곱 층과 Rectified Linear Unit(ReLU) 함수로 구성된 3개의 합성곱 블록(block)을 통과한 후 adaptive average pooling을 적용하였다.

실험 결과 모든 데이터 집합에서 TFT 방법에 사용했을 때 가장 정확한 예측이 되는 것을 확인할 수 있다. TFT 방법의 경우 10<sup>th</sup>, 50<sup>th</sup>, 90<sup>th</sup>로 분위수 예측하도록 설정했으며 각각

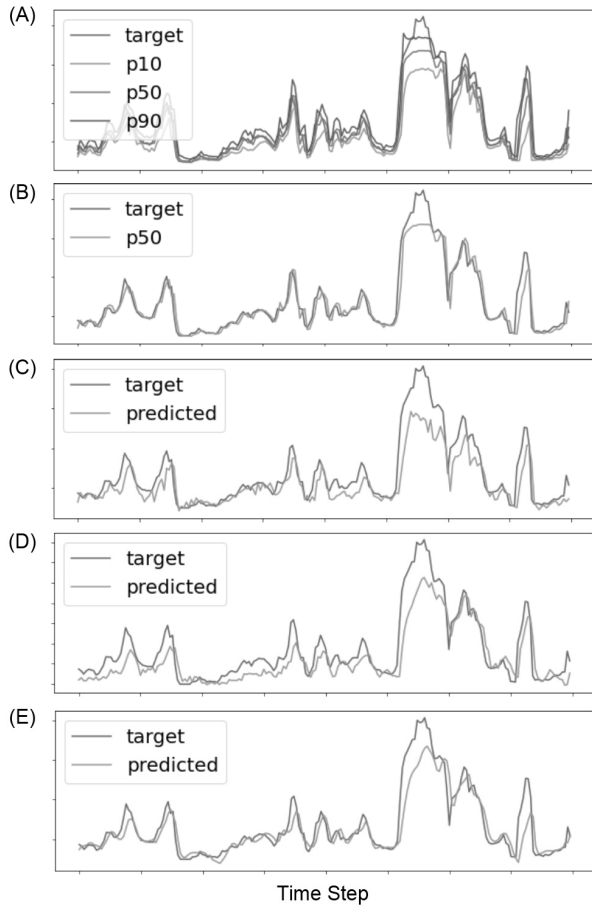


Fig. 2. Fine Dust Concentration Prediction Result from Beijing pm2.5 Dataset. (A) Prediction Result from TFT Method, (B) Prediction Result of 50<sup>th</sup> from TFT Method, (C) Prediction Result from Single LSTM, (D) Prediction Result from Stacked LSTM, (E) Prediction Result from 1D CNN

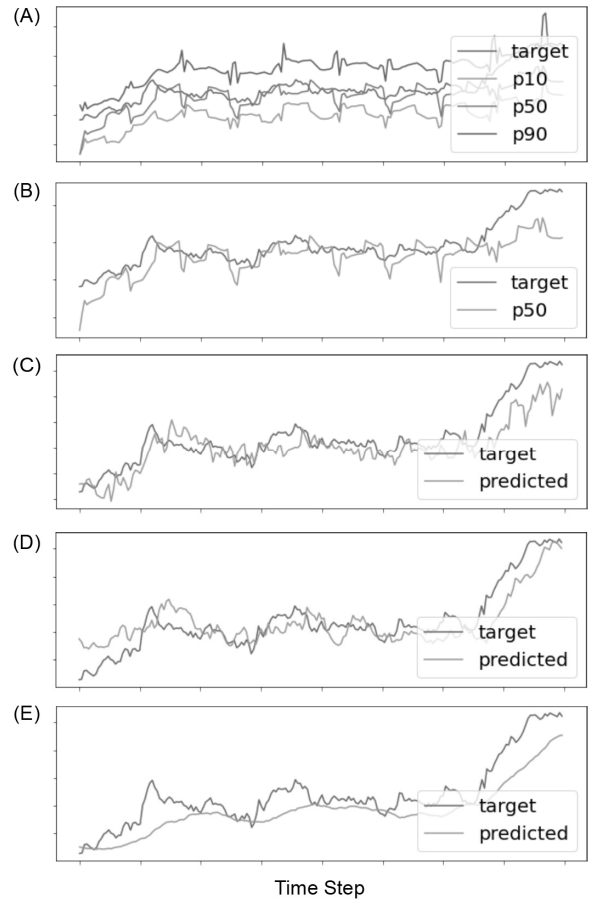


Fig. 3. Carriage Service Stock Prediction Result. (A) Prediction Result from TFT Method, (B) Prediction Result of 50<sup>th</sup> from TFT Method, (C) Prediction Result from Single LSTM, (D) Prediction Result from Stacked LSTM, (E) Prediction Result from 1D CNN

의 분위수에 대한 예측값을 Fig. 2A, Fig. 3A와 Fig. 4A에서 확인할 수 있다. 기존 모델과 비교하기 위한 값으로는 50<sup>th</sup>에서의 예측값을 사용할 수 있으며 명확한 비교를 위해 Fig. 2B, Fig. 3B와 Fig. 4B에 TFT 방법의 50<sup>th</sup> 예측 결과를 출력했다.

미세먼지 농도 데이터의 경우, Table 1에서 다층 LSTM 모델과 1D CNN 모델을 사용했을 때 Mean Squared Error(MSE) 값이 TFT 방법보다 상대적으로 조금 더 낮은 것을 확인할 수 있다. 하지만 예측한 결과를 출력한 Fig. 2A와 Fig. 2B를 통해서 TFT 방법을 사용한 실험 결과가 가장 정확했으며 Fig. 2C의 단층 LSTM 모델을 사용했을 때 예측 결과가 가장 정확하지 않은 것을 확인할 수 있다.

주식 데이터의 경우, Table 1에서 TFT 방법과 기존 방법들의 성능 차이가 매우 큰 것을 확인할 수 있다. 특히 TFT 방법과 다층 LSTM 모델의 MSE 값이 약 6.5배의 차이가 발생했다. 예측 결과를 출력한 Fig. 3을 통해서 1D CNN 모델은 LSTM 모델보다 상대적으로 부드럽게 예측한 것을 볼 수 있다. 주식 데이터 집합의 경우, 미세먼지 데이터보다 기존의

Table 1. Experimental Results

		Dataset		
		Beijing PM2.5 [10]	Carriage Service stock [9]	Electricity Transformer [11]
MSE Loss	TFT	0.1310	<b>0.0575</b>	<b>0.0462</b>
	Single LSTM	0.1282	0.2066	0.7785
	Stacked LSTM	0.2010	0.3765	0.6953
	1D CNN	<b>0.1116</b>	0.1583	0.4811

예측 방법인 LSTM이나 CNN 모델만을 사용했을 때 더 정확하게 예측되지 않은 것을 볼 수 있다. 이는 상대적으로 주식 데이터 집합의 입력값의 종류가 미세먼지 데이터와 비교해서 더 적었기 때문으로 해석된다.

마지막으로 전기 변압기 데이터를 사용해서 실험을 진행한 결과 TFT 방법을 사용했을 때 가장 정확하게 예측된 것을

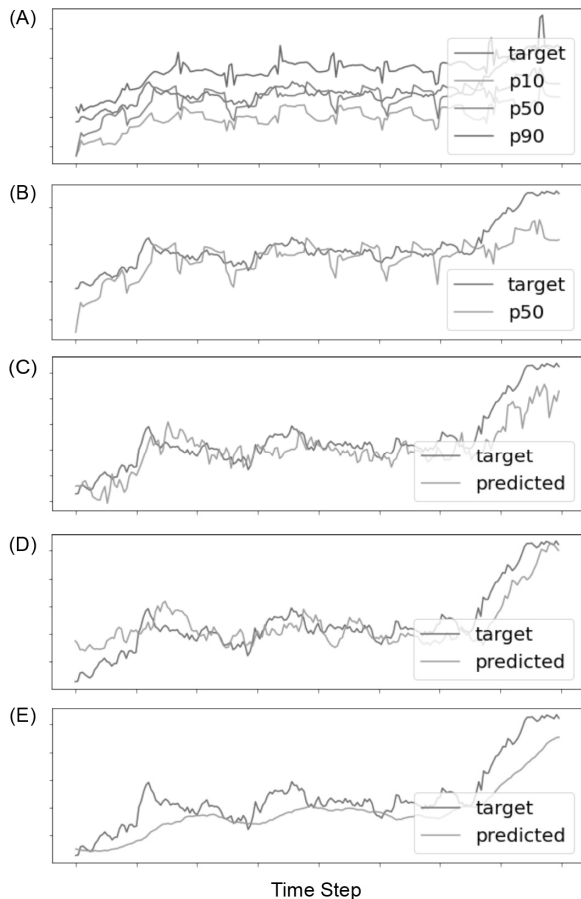


Fig. 4. Carriage Service Stock Prediction Result. (A) Prediction Result from TFT Method, (B) Prediction Result of 50<sup>th</sup> from TFT Method, (C) Prediction Result from Single LSTM, (D) Prediction Result from Stacked LSTM, (E) Prediction Result from 1D CNN

Fig. 4와 Table 1를 통해서 확인할 수 있다. 전기 변압기 데이터의 경우 다른 데이터 집합보다 기존 모델의 예측이 더욱 정확하지 않았다. 이는 정확한 전기 사용량 예측을 위해서는 연속적인 사용량에 대한 정보뿐만 아니라 장기적으로 계절과 함께 날씨 온도 정보 등을 함께 사용할 수 있어야 하기 때문이다. 하지만 LSTM이나 CNN 기반의 깊은 신경망 모델들은 관측된 입력값만을 사용한다는 한계로 제대로 된 예측이 이루어지지 않았다.

## 5. 결 론

실생활에서 사용되는 여러 데이터 종류가 시계열 형태를 이룬다는 점에서 정확한 다층 시계열 분석은 여러 분야에서 운영 효율성을 높일 수 있게 도와준다. 하지만 시계열 데이터는 다양한 종류의 정보를 포함하여 형성되기 때문에 이질적인 특성을 가진 데이터를 모두 반영해서 예측해야 정확한 예측 결과를 얻을 수 있다. 이를 위해서 본 논문에서는 TFT 방법과 기존의 심층학습 방법들을 실생활과 밀접한 관련이 있

는 데이터에 적용하여 예측 성능을 확인하고 성능을 정량적으로 비교하였다. 실험에는 미세먼지 데이터, 주식 데이터, 전기 변압기 데이터를 사용하였으며 실험 결과 TFT 방법이 다른 깊은 신경망 모델들보다 상대적으로 더 정확하게 예측하는 것을 확인하였다.

## References

- [1] B. Lim, S. O. Arik, N. Loeff, and T. Pfister, "Temporal fusion transformers for interpretable multi-horizon time series forecasting," *arXiv preprint arXiv:1912.09363*, 2019.
- [2] F. Luna-Perejon, M. J. Dominguez-Morales, and A. Civit-Balcells, "Wearable fall detector using recurrent neural networks," *Sensors*, Vol.19, No.22, pp.4885, 2019.
- [3] D. Kraft, K. Srinivasan, and G. Biebe, "Deep learning based fall detection algorithms for embedded systems, smartwatches, and IoT devices using accelerometers," *Technologies*, Vol.8, No.4, pp.72, 2020.
- [4] H. Hewamalage, C. Bergmeir, and K. Bandara, "Recurrent neural networks for time series forecasting: Current status and future directions," *International Journal of Forecasting*, Vol.37, No.1, pp.388-427, 2021.
- [5] D. Salinas, V Flunkert, and J. Gasthaus, "DeepAR: Probabilistic forecasting with autoregressive recurrent networks," *International Journal of Forecasting*, Vol.36, No.3, pp.1181-1191, 2020.
- [6] R. Wen, K. Torkkola, B. Narayanaswamy, and D. Madeka, "A multi-horizon quantile recurrent forecaster," *arXiv preprint arXiv:1711.11053*, 2017.
- [7] Y. Liu, C. Gong, L. Yang, and Y. Chen, "DSTP-RNN: A dual-stage two-phase attention-based recurrent neural network for long-term and multivariate time series prediction," *Expert Systems with Applications*, Vol.143, pp.113082, 2020.
- [8] C. Fan, Y. Zhang, Y. Pan, X. Li, C. Zhang, and R. Yuan, "Multi-horizon time series forecasting with temporal attention learning," *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp.2527-2535, 2019.
- [9] C. Favorita, Corporacion favorita grocery sales forecasting competition, 2009.
- [10] Yahoo Finance, Carriage Services, Inc., 2021.
- [11] X. Liang, T. Zou, B. Guo, and S. Li, "Assessing Beijing's PM2.5 pollution: Severity, weather impact," *APEC and Winter heating, Proceedings of Royal Society*, Vol.471, No.2182, pp.20150257, 2015.
- [12] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, and H. Xiong, "Informer: Beyond efficient transformer for long sequence time-series forecasting," *Proceedings of AAAI*, 2021.



**김 인 경**

<https://orcid.org/0000-0002-6373-7646>  
e-mail : kim\_jenn@kookmin.ac.kr  
2020년 국민대학교 정치외교학부(학사)  
2020년~현 재 국민대학교 컴퓨터공학과  
석사과정  
관심분야: Artificial Intelligence(AI) &  
Time Series Data Analysis



**이 재 구**

<https://orcid.org/0000-0002-5947-5487>  
e-mail : jaekoo@kookmin.ac.kr  
2018년 서울대학교 전기컴퓨터공학부(박사)  
2011년~2013년 LG전자 CTO부문  
주임연구원  
2018년 SK텔레콤 ICT기술원 매니저  
2018년~현 재 국민대학교 SW학부 조교수  
관심분야: Artificial Intelligence(AI), Machine Learning,  
Deep Learning and Data Science



**김 대 희**

<https://orcid.org/0000-0001-9676-9604>  
e-mail : frshield2@kookmin.ac.kr  
2020년 국민대학교 소프트웨어전공(학사)  
2020년~현 재 국민대학교 컴퓨터공학과  
석사과정  
관심분야: Artificial Intelligence(AI) &  
Domain Generalization