

단기 시계열 데이터에 대한 장기간 다단계 예측을 위한 학습 방법 연구

김동현, 도주성, 정호영*
경북대학교

kdh1477@knu.ac.kr, jsdo@knu.ac.kr, *hoyjung@knu.ac.kr

A study on learning method for multi-step long-term forecasting using short-term time series data

Dong-Hyun Kim, Ju-Seong Do, Ho-Young Jung*
Kyungpook National Univ.

요 약

본 논문은 기존 시계열 데이터 예측의 문제점을 해결하기 위해 단기간 (1 일)의 데이터를 활용하여 장기간 (1 달)의 데이터를 예측하는 연구에 대해 소개한다. 기존의 시계열 데이터 예측에서 단기간에 대해서는 준수한 성능을 보이지만, 장기간은 거의 예측하지 못하는 문제점이 있다. 이를 해결하기 위해, 본 논문에서는 단기간 예측 성능이 우수하다는 점을 활용하여 이를 연속으로 수행함으로써 보다 나은 장기간 예측이 가능한 방법을 제안한다. 장기간 예측에서 제안된 방법이 기존의 예측 방법보다 목표값을 더 잘 추적하는 경향을 보인다. 본 논문은 제안된 학습 방법이 단기간의 시계열 데이터를 활용한 데이터 예측 모델 개발에 응용될 수 있을 것으로 기대한다.

I. 서론

최근 인공지능을 기반으로 시계열 데이터 예측을 위한 기술들이 개발되고 있다. 기존의 시계열 데이터 예측에 사용된 모델은 순환 신경망인 RNN (Recurrent Neural Network)이다. RNN의 구조를 가지고 있는 대표적인 모델에는 LSTM[1] (Long Short-Term Memory), GRU[2] (Gated Recurrent Units)등이 있다. 위 모델들은 장기간 데이터 학습에 어려움이 있고 학습 시간이 오래 걸리는 등의 단점이 있다. 이후 등장한 Transformer[3] 모델은 시계열 데이터의 병렬처리로 RNN의 학습 시간 문제를 해결하였다.

본 논문에서는 Transformer 모델 기반의 Informer[4] 모델을 LSTF (Long Sequence time-series forecasting) 태스크에 활용한다. Informer 모델은 기존의 시계열 데이터 예측 모델들보다 더 긴 sequence를 예측하는 성과가 있지만, 매우 긴 sequence를 예측하는 데는 기존의 모델과 비슷한 성능을 보인다.

위의 문제를 해결하기 위해, 본 논문에서는 단기간 예측이 우수하다는 점에 주목하여 장기간을 예측하는데 단기간 예측을 그 다음 단기간 예측에 활용하여 연속적으로 추론하는 방법을 시도한다. 이러한 장기간 다단계 예측을 학습에도 적용하고 개선된 장기간 예측의 결과를 보여준다.

II. 본론

본 장에서는 실험에 사용된 데이터셋 (ETDataset), 모델 (Informer), 장기간 다단계 예측을 위한 학습 방법에 대해 소개한다.

II - I . ETDataset (Electricity Transformer Dataset)

본 논문에서 LSTF 태스크에 사용한 학습데이터는 ETDataset으로 실제 중국에서 2년간 2곳의 Electricity Transformers에서 수집된 데이터이다. ETDataset은 1시간 단위인 {ETTh1, ETTh2}와 15분 단위인 {ETTm1, ETTm2}로 나뉘어 있다. 각 데이터의 변수는 총 7개로 부하특성인 HUFL (High UseFul Load), HULL (High UseLess Load), MUFL (Middle UseFul Load), MULL (Middle UseLess Load), LUFL (Low UseFul Load), LULL (Low UseLess Load)와 목표값인 OT (Oil Temperature)로 구성되어 있다.¹

II - II . Informer

Informer 모델은 기존의 Transformer 모델에서 사용하던 self-attention을 ProbSparse self-attention으로 바꿔 연산량을 줄여 좀 더 빠른 학습이 가능하게 한다. Transformer의 self-attention은 key와 query의 내적값을 모두 이용하여 연산을 수행하는데, Informer의 ProbSparse self-attention은 key와

¹ <https://github.com/zhouhaoyi/ETDataset>

query의 내적값 중 큰 값만을 가려내어 연산을 수행한다. 또한, Informer는 Transformer와 달리 convolution layer를 추가하여 데이터의 특징을 좀 더 잘 찾아내도록 한다. 결과적으로, Transformer가 N^2 만큼의 시간과 메모리를 소비하며 학습할 때, Informer는 $N\log(N)$ 만큼의 시간과 메모리를 소비하며 학습한다. (N 은 sequence의 길이를 의미)

II-III. 장기간 다단계 예측을 위한 학습 방법

본 논문은 단기간 미래 예측 모델을 연속적으로 적용하여 장기간 미래 예측시스템을 위한 WFO (Walk-Forward Optimization) 기반의 학습 방법을 제안한다. 이를 위해 Informer의 criterion function을 개선한다. 식 (1)의 L_1 은 기존의 Informer에 사용된 loss이다. 즉, 실제 과거 1일의 데이터로 미래 1일의 데이터를 찾는 모델의 loss이다. WFO를 위한 L_2 는 임의의 1일 데이터에 대한 미래 1일 예측값을 다시 입력으로 하여 다음 미래 1일 데이터를 예측하는 criterion이다.

$$\text{Loss} = \lambda L_1 + (1 - \lambda)L_2 \quad (1)$$

실험에 사용된 데이터셋은 ETDataset의 ETTh1의 첫 번째 변수인 HUFL이다. 실험은 NVIDIA Quadro RTX 24GB GPU로 진행되었으며, 학습률은 $1e-4$, batch size는 24, λ 는 0.5, 옵티마이저는 Adam을 사용하였다. 평가는 각 모델에 대해 30번씩 수행되었으며 실험 결과에 기술된 MSE (Mean Square Error)는 30번 실험의 평균치이다.

II-IV. 실험 결과

제안하는 학습 방법의 성능을 평가하기 위해, 기존의 방법으로 학습된 모델의 추론 결과와 제안된 방법으로 학습된 모델의 추론 결과를 비교한다.

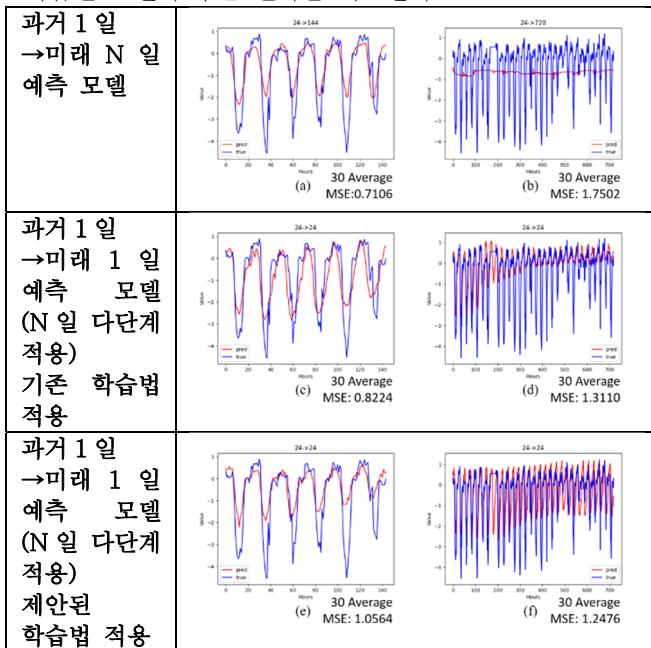


그림 1: 추론 결과 비교.

그림 1의 (a, b)는 sequence 길이에 따른 Informer 성능 비교 결과이다. 그림 1의 (a)는 1일치 데이터로 미래의 6일치 데이터를 예측한 결과이고, (b)는 1일치 데이터로 미래의 30일치 데이터를 예측한 결과이다. 그림 1의 (c, d)는 1일치 데이터로 다음 1일치 데이터를 예측하고 예측한 데이터로 그 다음 1일치 데이터를 연속적으로 예측한 결과이다. 그림 1의 (e, f)는 장기간 다단계 예측 방법을 학습에도 적용하고 새롭게 학습된 모델로 장기간을 다단계 예측하는 LSTF 결과이다. 즉,

그림 1의 (a, c, e)는 총 6일의 기간을 예측한 것이고, 그림 1의 (b, d, f)는 총 30일의 기간을 예측한 것이다.

6일의 기간을 예측할 때는 그림 1의 (a, c, e)에서 볼 수 있듯이 기존의 모델로 한 번에 6일을 예측한 결과와 장기간 다단계 예측 모델로 6일 동안 예측한 결과가 유사하다. 30일의 기간을 예측할 때는 그림 1의 (b)와 같이 기존의 모델로 한 번에 30일을 예측한 결과는 평균값을 추적하는 경향이 있지만, 장기간 다단계 예측 모델로 30일 동안 예측한 결과는 그림 1의 (d, f)에서 보이듯이 변화를 지속적으로 추적하고 있음을 확인할 수 있다.

III. 결론

본 논문에서는 시계열 데이터의 장기간 예측에 집중하였다. 이를 위해 기존 연구와 달리 시계열 데이터의 연속적인 추론을 통해서 LSTF 태스크를 수행하는 다양한 실험을 하였다. 30일의 장기간을 예측할 때 기존의 예측 모델은 평균값을 추적하는 경향이 있다. 반면, 제안된 장기간 다단계 예측법은 변화를 따라감과 동시에 이를 학습에 적용하면 오차 또한 줄어드는 것을 확인할 수 있다. 이러한 결과는 시간의 흐름을 가진 모든 데이터에 충분히 적용될 수 있을 뿐만 아니라 주가 예측, 설비 예지보전, 날씨 예측 등에도 적용될 수 있을 것으로 기대된다.

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT 연구센터지원사업의 연구결과로 수행되었음 (IITP-2021-2020-0-01808*)

참 고 문 헌

- [1] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9.8 (1997): 1735-1780.
- [2] Cho, Kyunghyun, et al. "Learning phrase representations using RNN encoder-decoder for statistical machine translation." arXiv preprint arXiv:1406.1078 (2014).
- [3] Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems. 2017.
- [4] Zhou, Haoyi, et al. "Informer: Beyond efficient transformer for long sequence time-series forecasting." Proceedings of AAAI. 2021.
- [5] Chung, Junyoung, et al. "Empirical evaluation of gated recurrent neural networks on sequence modeling." arXiv preprint arXiv:1412.3555 (2014).