

실생활의 불안정한 변동 요인을 고려한 효율적인 장기 시계열 예측 모델 연구

GD - MLP (A study of an efficient long-term time series forecasting model
considering unstable fluctuation factors in real life)

설근형 (Seol, Geun Hyung)

요 지 (Abstract)

시계열 속 긴 시점의 변동을 포착하는 장기 시계열 예측(이하 LSTF)은 의사결정의 범위를 넓혀 실생활의 응용에 큰 효용을 가져다 줄 수 있으나 이는 여전히 어려운 과제로 남아있다. 왜냐하면 시계열 예측 문제가 응용되는 여러 도메인은 추세 변동과 계절 변동 등 각각이 복잡한 변동들로 얽혀 있으며, 예측시점인 미래는 그러한 변동들이 과거와 전혀 다른 양상을 보일 수 있는 매우 불안정한 환경이기 때문이다. 그러나 그동안 트랜스포머가 주도한 LSTF 연구들은 이러한 실생활의 응용 상황을 고려하지 않은 채, 과거의 변동을 정교하게 포착할 수 있는 복잡한 구조를 디자인하는 것에만 몰두해왔다. 그리고 그 결과 최근의 연구들은 그들의 한계점을 분명히 드러냄과 동시에 더 가볍고 유연한 모델이 트랜스포머의 예측 성능을 압도한다는 결과를 보여주고 있다. 이에 본 연구에서는 트랜스포머 모델들의 한계점과 최근의 연구 결과를 바탕으로 실생활의 응용에 가장 적합한 모델을 연구하였다. 본 연구에서 제안하는 GD-MLP(Gated Decomposition MLP)는 실생활의 응용을 최우선으로 고려한 모델로, 요소 분해와 두 가지 게이트(Gate)를 통해 도메인에 따른 지배적인 변동은 물론 불안정한 변동요인에 유연하게 대처할 수 있는 구조를 가지고 있다. 그 결과 팬데믹(pandemic)으로 인해 시계열의 추세가 매우 불안정했던 2020년 매출 데이터에서 높은 정확도를 입증하였으며, 그동안 LSTF 모델들이 실험한 벤치마크 데이터에서도 높은 정확도를 확인하였다. (Under Review)

Long sequence time-series forecasting (LSTF), which captures long-term fluctuations in time series, can bring great utility to real-life applications by broadening the scope of decision-making, but it remains a difficult task. This is because the various domains to which time series forecasting problems are applied are entangled with complex fluctuations such as trend fluctuations and seasonal fluctuations, and the future, which is the forecasting point, is a very unstable environment in which such fluctuations can show completely different aspects from the past. However, LSTF research led by transformers has been focused only on designing complex structures that can accurately capture past fluctuations without considering these real-life applications. And as a result, several recent studies clearly reveal their limitations and show results that lighter and more flexible models overwhelm the forecasting performance of transformers. Therefore, based on the limitations of transformer models and recent research results, I study the most suitable model for real-life applications. The model, GD-MLP (Gated Decomposition MLP) proposed in this study is a model that prioritizes real-life applications and can flexibly capture unstable fluctuations as well as dominant fluctuations according to domains through series-decomposition and two gates. As a result, this model proved the high accuracy in benchmark data tested by LSTF models so far.

1. 서론

시계열 예측 모델은 날씨, 금융 공정, 비즈니스 등 실생활의 여러 도메인(domain)에서 유용하게 이용되고 있다. 특히나 보통 100 시점 이상을 예측하는 문제 (1)를 말하는 장기 시계열 예측(이하 LSTF)은 다양한 의사결정을 지원하여 분석의 활용성을 크게 높일 수 있다. 그런데 그러한 LSTF를 효과적으로 수행하는 모델을 디자인하는 일은 여전히 어려운 과제로 남아있다. 왜냐하면 현실 속 시계열의 패턴은 각 도메인에 따라 복잡하고 특징적인 변동들로 이루어져 있으며, 예측의 대상인 미래 시점은 때에 따라 과거의 시계열과 전혀 다른 양상을 보일 수 있는 매우 불안정한 환경이기 때문이다. 따라서 실생활의 응용에 가장 적합한 LSTF 모델은 지엽적인 변동에 과대적합되는 것을 경계해야 하며, 도메인의 특징적인 변동과 미래의 위험성을 잘 반영할 수 있는 유연하고 효율적인 구조를 가져야 한다.

그런데, 최근 LSTF의 주요 연구대상인 트랜스포머는 위와 같은 실생활의 응용에 필요한 조건들을 충족하지 못하고 있다. 트랜스포머는 기본적으로 모든 포인트와의 유사성을 추출하는 어텐션 메커니즘을 통해 시간적 종속성(temporal dependency)을 먼 시점까지 효과적으로 추출 (2)하여 LSTF에서 높은 정확도를 보이기도 했지만, 복잡한 구조와 높은 연산 비용에 의해 효율성이 매우 떨어지는 문제가 있다. 이에 많은 연구들이 다양한 방법으로 트랜스포머의 비효율성을 개선함과 동시에, 예측 정확도를 높이기 위해 자기상관 매커니즘과 시계열 변동의 요소 분해 등 시계열 분석의 특징적인 요소들을 추가한 변형 모델들을 제안하기도 하였다. 그러나 위와 같은 연구로 등장한 트랜스포머 변형 모델¹

또한 효율성과 효과성의 측면에서 한계점이 드러나고 있다. 최근의 연구 (3)에서는 다양한 실험을 진행하여 트랜스포머 변형 모델들의 LSTF 능력을 평가하였는데, 해당 실험²에서는 연구 논문들의 주장과 달리 트랜스포머 변형 모델들의 효율성이 이론만큼 크게 개선되지 않았으며, 그들의 지나치게 복잡한 구조가 과대적합 문제를 야기하여 시계열의 변동을 효과적으로 추출하지 못하고 있음을 밝혀냈다. 반대로 이 실험에서는 트랜스포머와 대조되는 단순한 선형 모델을 실험의 기준선 모델³로 활용하였는데 그 모델은 실험에서 매우 우월한 효율성과 함께 시계열의 변동을 효과적으로 포착하는 모습을 보여주었다. 즉, 해당 실험 결과는 트랜스포머의 한계점을 확인한 것과 더불어, 더 가벼운 모델이 LSTF에 있어서 트랜스포머의 좋은 대안이 될 수 있다는 가능성을 시사한 것이다.

이에 본 연구에서는 앞서 언급한 모든 아이디어를 종합하여 LSTF에 가장 적합한 모델을 연구하였다. 본 연구에서 가장 우선으로 고려한 것은 바로 **실생활의 응용을 위한 유연성과 효율성의 확보**이다. 이를 위해 트랜스포머의 가열된 연구에서 벗어나 상대적으로 가벼운 **MLP 구조**를 기반으로 한 GD-MLP(Gated Decomposition MLP) 모델을 디자인하였다. 이 모델은 시계열 요소 분해를 통해 추세 변동과 계절 변동을 포함한 나머지 변동을 각각 MLP에 적용하는데, 이때 게이트를 배치하여 변동들을 선택적으로 학습하는 것이 특징이다.

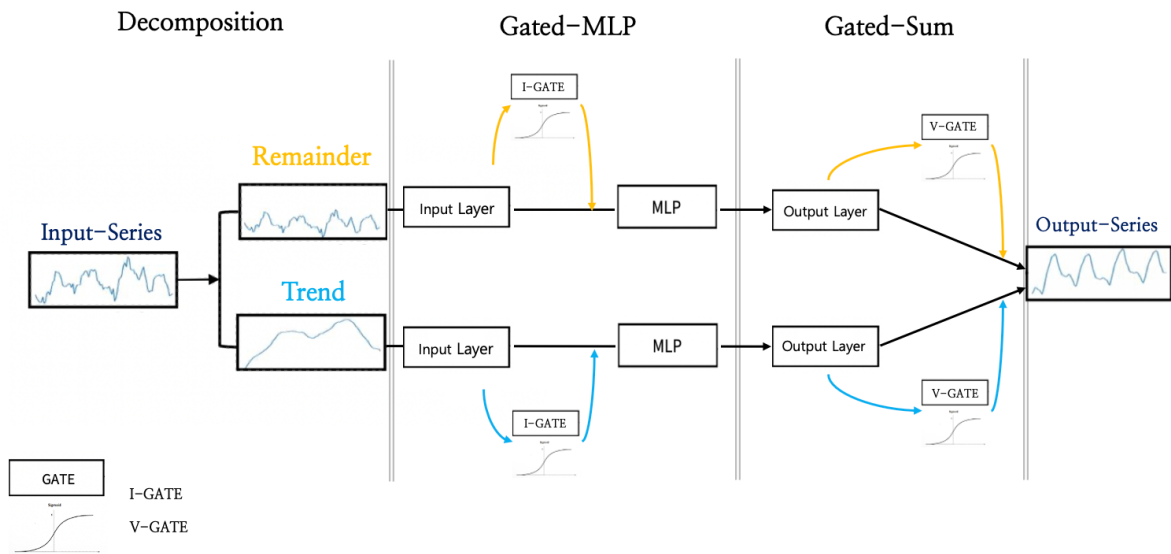
마지막으로 이 모델의 유연성과 효율성을 확인하기 위해 실생활의 응용 상황을 가정하여 실험을 디자인하였다. 이에 실험에 쓰인 데이터는 실생활의 비즈니스 도메인에서 여러 변동에 의해 추세가 매우 탄력적인 매출 데이터와 팬데믹 상황으로 인해 추세가 매우 불안정했던 2020년 시점을 선정하였다.

모델 아키텍처

¹ Informer, Autoformer, FED-former 등 트랜스포머를 기반으로 각자 다양한 아이디어를 가미한 모델이 제안되었다.

² (3)의 6p~8p에서 진행한 실험들

³ (3)에서는 기준선 모델로 단순히 하나의 레이어를 배치한 D-linear, N-linear 모델들을 실험에 투입하였다. D-linear 모델은 요소 분해를 적용한 구조이며, N-linear는 분포의 변화를 포착하기 위해 가장 최근값으로 정규화를 진행한 모델이다.



참고 문헌

1. "Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. "Haoyi ZhouZhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, Wancai ZhangShanghang. : AAAI, 2021. AAAI.
2. "Time Series Forecasting With Deep Learning: A Survey. "Zohren1Lim1 and StefanBryan.
3. "Are Transformers Effective for Time Series Forecasting?(2022). "Ailing ZengChen, Lei Zhang, Qiang XuMuxi. 2022.
4. "Transformers in Time Series: A Survey. "Qingsong WenZhou, Chaoli Zhang, Weiqi Chen, Ziqing Ma, Junchi Yan, Liang SunTian. : IJCAI , 2022.
5. "Autoformer: Decomposition Transformers with Auto-Correlation for Long-Term Series Forecasting. "Haixu WuXu, Jianmin Wang, Mingsheng LongJiehui.: NeurIPS, 2021.
6. "Reformer THE EFFICIENT TRANSFORMER. "Nikita KitaevKaiser, Anselm LevskayaŁukasz. ICLR , 2020.
7. "Less Is More: Fast Multivariate Time Series Forecasting with Light Sampling-oriented MLP Structures. "Tianping ZhangZhang, Wei Cao, Jiang Bian, Xiaohan Yi, Shun Zheng, Jian LiYizhuo. cs.LG.