

Transformer-XGBoost 모델을 활용한 계통 한계 가격 핵심 영향 요인 분석

최우주, 신한솔*, 김선혁
국립공주대학교

woozoo1007@smail.kongju.ac.kr, hansol.shin@etri.re.kr*, seonh@kongju.ac.kr

Analyzing the Key Drivers of System Marginal Price using a Transformer-XGBoost Hybrid Model

Woo Ju Choi, Han Sol Shin*, Seon Hyeog Kim
Kongju National Univ, *Electronics and Telecommunications Research Institute (ETRI)

요약

본 논문은 계통 한계 가격(System Marginal Price) 변동에 영향을 미치는 핵심 외부 요인을 규명하는 것을 목표로 한다. 이를 위해, 과거 SMP 데이터에 의존하지 않고 외부 시장 및 기상 정보만을 활용하는 Transformer-XGBoost 하이브리드 모델을 제안한다. 제안 모델은 시계열 데이터의 동적 특징을 효과적으로 추출하여 RMSE 12.37 을 기록했으며, XGBoost 단일 모델의 14.28 대비 높은 예측 정확도를 보이는 수치이다. 이 모델의 특징 중요도 분석 결과, 24 시간 전 최대 전력 수요, 요일, 시간 정보가 SMP 예측에 핵심적인 요인임을 확인하였다. 이는 제안된 모델이 SMP의 복잡한 변동성을 효과적으로 예측할 뿐만 아니라, SMP를 결정하는 주요 외부 요인을 성공적으로 식별할 수 있음을 보여준다.

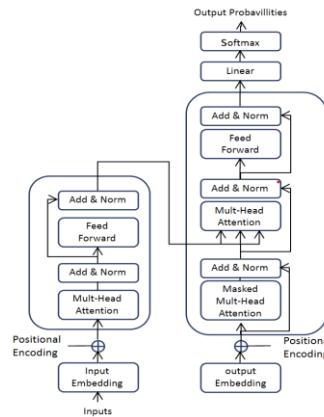
I. 서론

계통 한계 가격(System Marginal Price)은 전력시장의 핵심 지표이지만, 전력 수요, LNG 가격 등 다양한 외부 요인에 의해 복잡하게 변동하여 어떤 요인이 얼마나 영향을 미치는지 분석하기 어렵다. 기존 연구들은 예측 성능을 높이기 위해 과거 SMP 데이터를 주요 입력 변수로 사용하였으나, 이는 모델이 과거 SMP 값에 과도하게 의존하게 만들어, SMP 변동을 유발하는 외부 요인들의 영향력을 명확히 파악하기 어렵게 한다. 본 연구의 주된 목적은 SMP에 영향을 미치는 핵심 외부 요인을 규명하는 것으로, 과거 SMP 데이터를 입력에서 제외하고 오직 외부 시장 및 기상 정보만을 활용하는 새로운 접근법을 제시한다. 요인 분석은 높은 예측 정확도를 가진 모델을 기반으로 할 때 신뢰성을 갖기 때문에, 시계열 데이터의 시간적 맥락 추론에 탁월한 Transformer 와 강력한 예측 성능을 보이는 XGBoost 를 결합한 하이브리드 모델을 제안한다. 이를 통해 SMP에 영향을 미치는 주요 외부 요인을 명확히 식별하고자 한다.

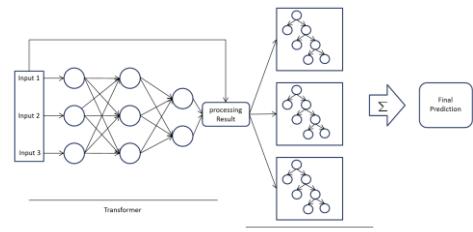
II. 본론

2.1 Transformer-XGBoost 기반 SMP 예측 모델

Transformer 는 셀프 어텐션(Self-Attention) 메커니즘을 통해 시계열 데이터의 복잡한 시간적 패턴과 변수 간 상호 작용을 학습한다. Transformer 는 인코더와 디코더로 구성되며, 인코더는 데이터를 이해하고 문맥을 압축하며, 디코더는 이 문맥을 바탕으로 결과를 생성한다. 이러한 Transformer 의 구조는 <그림 1>과 같으며, Transformer 가 인코더에서 생성한 문맥 정보와 이전 차신의 출력값을 참고하여 디코더가 효과적으로 학습하는데 기여한다.



<그림 1> Transformer 내부 모델 [1]



<그림 2> Transformer-XGBoost 내부 모델 [2][3]

XGBoost 모델은 <그림 2> 우측과 같이 여러 개의 결정 트리를 결합하여 예측 성능을 극대화하는 그레디언트 부스팅 알고리즘이다. 단일 트리의 예측 성능 저하 문제를 해결하기 위해 이전 트리의 오차를 다음 트리가 보완해 나가는 방식으로 학습하여 높은 예측 정확도와 빠른 속도를 가진다.

본 연구에서는 시계열 데이터의 예측 정확도를 높이기 위해서 Transformer 와 XGBoost 를 결합한 Transformer-XGBoost 모델을 사용하였다. 모델의 구조는 <그림 2>와 같이 특징 추출을 담당하는 Transformer

와 최종 예측을 수행하는 XGBoost로 구성된다. 먼저, Transformer는 메커니즘을 통해 원본 데이터에 내재된 동적 맥락 정보를 함축하여 시간적 특징을 추출하고, XGBoost는 부스팅 기법을 통해 최종 예측값을 산출한다. 이러한 2 단계 하이브리드 구조를 통해 정교한 패턴 인식과 예측 성능을 향상시켰다.

먼저, Transformer는 셀프 어텐션 메커니즘을 통해서 학습하여 원본 데이터에 내재된 동적 맥락 정보를 함축해 시간적 특징을 추출한다. 이후 XGBoost는 부스팅 기법을 통해 최종 예측값을 산출한다. 이러한 2 단계 하이브리드 구조를 통해 정교한 패턴 인식과 강력한 예측 성능 정확성을 극대화하였다.

2.2 모델 학습

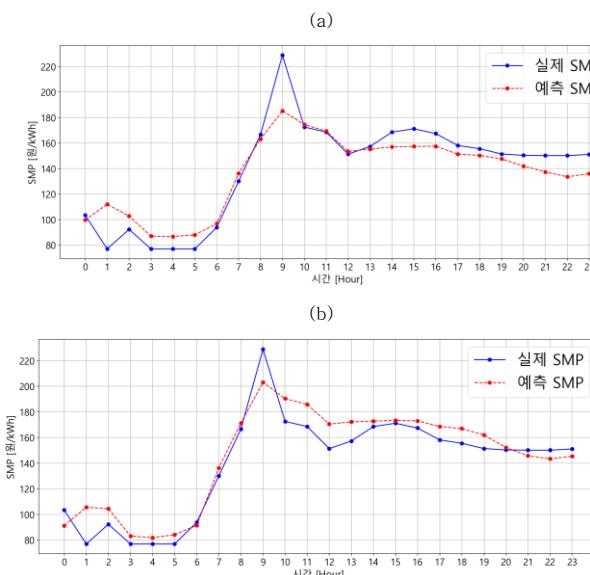
본 연구에서 사용된 데이터는 국내외 공공기관에서 제공하는 시계열 데이터로 2018년 1월 1일부터 2024년 12월 31까지의 데이터를 대상으로 하였다. 상세내용은 <표 1>과 같다.

<표 1> 실험 사용 데이터

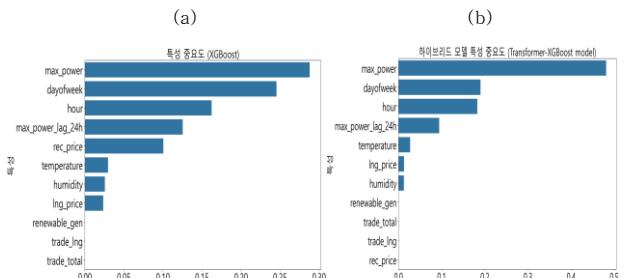
| 데이터 종류 | 제공기관 | 내용 | 변수 |
|----------|--------------------|--------------------------|--------------------------------|
| SMP | 전력 거래소(KPX) | 시간별 유통 한계 가격 | x |
| LNG 가격 | 미국 EIA | 천연가스 현물 가격 | lng_price |
| 최대 전력 수요 | 전력 통계정보 시스템(EPSIS) | 최대 전력 수요 하루 전 전력수요 | max_power max_power_lag_24h |
| 기상 데이터 | 기상청 | 기온, 습도 | temperature, humidity |
| 발전량 | 전력 통계정보 시스템(EPSIS) | 1년 신재생 발전량 | renewable_gen |
| 전력 거래량 | 전력 통계정보 시스템(EPSIS) | 달 별 가스 거래량 달 별 전력 거래량 | trade_lng trade_total |
| REC 가격 | 전력 거래소(KPX) | 1년 평균값 | rec_price |

SMP의 데이터의 경우 결측값은 24시간 전의 동일 시간대 값으로 대체하여 데이터의 주기성을 보존하고, 모델 성능 향상을 위해 날짜 정보에서 요일, 시간 등의 주기적 특징을 추출하였으며, 24시간 전 최대 전력 수요와 같은 변수를 생성하여 과거 정보를 활용할 수 있도록 설계하였다.

2.3 결과 및 성능



<그림 3> 실제 SMP 값을 각 모델로 예측 결과(XGBoost, T-X 모델) 하루의 실제 SMP 값과 각 모델의 예측 결과를 시각적으로 비교한 것이 <그림 3>이다.



<그림 4>모델 별 중요도 분석(XGBoost, Transformer-XGBoost)

<표 2>모델 별 오차(RMSE) 비교

| 모델(Model) | RMSE |
|---------------------------|-------|
| XGBoost(비교모델) | 14.28 |
| Transformer-XGBoost(제안모델) | 12.37 |

(a)의 XGBoost 단일 모델은 <표 2>에서 RMSE 가 14.28 을 기록한 반면, (b)의 하이브리드 모델은 RMSE 12.37 을 기록하며 실제값을 더 정밀하게 추종하는 것을 확인할 수 있다.

<그림 4>는 제안 모델이 예측을 수행할 때 어떤 외부 요인을 중요하게 고려했는지 분석한 결과이다. 두 모델 모두 전력량(24시간 전 최대 수요), dayofweek(요일), hour(시간)에 큰 영향을 받았으며, 그 외 데이터들은 상대적으로 영향이 적거나 모델에 따라 중요도가 다르게 나타나는 것을 확인할 수 있다.

III. 결론

본 논문에서는 과거 SMP 데이터에 의존하지 않고 외부 요인 만으로 SMP 변동을 분석하기 위해 Transformer-XGBoost 하이브리드 모델을 제안하였다. 모델의 중요도 분석을 통해서 24시간 전 최대 전력 수요, 요일, 시간이 SMP를 결정하는 가장 지배적인 외부 요인임을 실증적으로 확인하였다. 이는 과거 SMP 값에 의존하지 않고도 외부 요인의 순수한 영향력을 효과적으로 분석할 수 있음을 의미한다. 나아가 본 연구가 지능형 에너지 예측 시스템 구축의 기초가 되어, 보다 효율적인 에너지 관리 체계에 기여할 것을 기대한다.

ACKNOWLEDGMENT

본 연구는 산업통상자원부(MOTIE)와 한국에너지기술평가원(KETEP)의 지원 (No. RS- 2023-00237018), 그리고 국립공주대학교 학술연구지원사업의 의하여 연구되었음.

참 고 문 헌

- [1] A. Vaswani *et al.*, "Attention is all you need," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017.
- [2] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2016, pp. 785-794.
- [3] Y. Gu, Z. Zhang, J. Li, Y. Li, and H. Zhang, "A novel hybrid model of CNN and XGBoost for groundwater level forecasting," *Complex & Intelligent Systems*, vol. 9, pp. 1235-1249, 2023.