

수치 모델을 활용한 지식 증류 기반 기상 예측 딥러닝 모델

유선희¹, 정은성²

¹홍익대학교 소프트웨어융합학과 학부생

²홍익대학교 소프트웨어융합학과 부교수

dup08063@naver.com, ejung@hongik.ac.kr

Deep Learning Model for Weather Forecast based on Knowledge Distillation using Numerical Simulation Model

요 약

딥러닝에서 지식 증류 기법은 큰 모델의 지식을 작은 모델로 전달하여 작은 모델의 성능을 개선하는 방식이다. 지식 증류 기법은 모델 경량화, 학습 속도 향상, 학습 정확도 향상 등에 활용될 수 있는데, 교사 모델이라 불리는 큰 모델은 일반적으로 학습된 딥러닝 모델을 사용한다. 본 연구에서는 학습된 딥러닝 모델 대신에 수치 기반 시뮬레이션 모델을 사용함으로써 어떠한 효과가 있는지 검증하였으며, 수치 모델을 활용한 기상 예측 모델에서의 지식 증류는 기존 단독 딥러닝 모델 학습 대비 더 작은 학습 횟수(epoch)에서도 동일한 에러 수준(RMSE)까지 도달하여, 학습 속도 측면에서 이득이 있음을 확인하였다.

1. 서론

4 차 산업 혁명은 인공지능과 빅데이터의 발전은 현대 사회의 여러 분야에서 혁신적인 변화를 가져왔다. 이러한 기술의 발전은 기상 예측 분야에도 많은 영향을 미치고 있다. 기존의 기상예측을 위해서는 전 지구예보모델, 국지모델 등의 수치예보모델을 기반으로 하였지만, 인공지능이 발전함에 따라 예측 모델에도 많은 발전이 이루어졌고, 이는 정확도의 향상을 가져왔다. 하지만 이러한 새로운 기술이 도입됨에 따라 기상예측 모델은 무거워졌고, 고성능의 컴퓨팅 장치를 요구하고 있다. 이러한 문제점을 해결하기 위해 본 연구에서는 딥러닝과 같은 기술을 활용하여 기상 예측 모델의 정확도를 높이고 모델을 경량화 하기 위한 연구를 수행하였다. 특히, 이 연구에서는 모델 경량화 방법으로 지식 증류(Knowledge Distillation) 기술을 적용하여 기상예측 모델의 성능을 개선하고자 한다.

2. 지식 증류(Knowledge Distillation)

지식 증류 기술은 큰 모델(교사 모델)의 지식을 작은 모델(학생 모델)로 전달하여 작은 모델의 성능을 개선하는 방식이다 [1]. 지식의 이전을 통해 작은 모

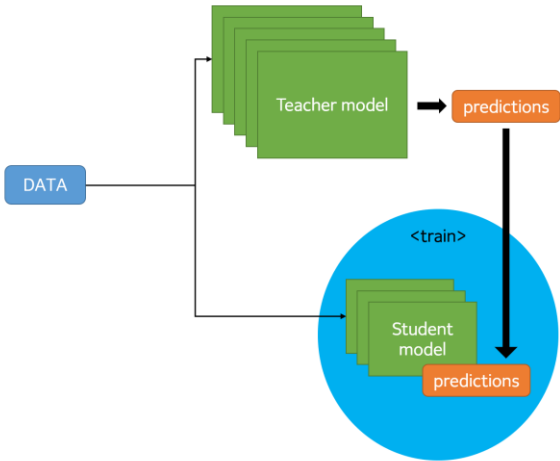
델은 큰 모델의 성능을 모방하도록 학습되어 더 빠르게 예측 모델을 구현할 수 있다. 또한, 지식 증류는 작은 모델의 불필요한 파라미터를 배제하여 모델의 경량화를 실현할 수 있게 하여 더 효과적으로 작동할 수 있게 한다. 일반적으로 교사 모델은 이미 학습된 모델을 활용하며, 예를 들어 이미지 분류 문제의 경우 교사 모델로 ResNet 모델을 활용하는 식이다.

3. LATM AutoEncoder

모델은 이전 연구 [2]에서 LSTM 활용 모델 중 가장 좋은 성능을 보인 LSTM AutoEncoder 를 사용하였다. LSTM AutoEncoder 는 AutoEncoder 의 구조에 LSTM layer 를 적용한 모델이다. Encoder 와 Decoder 에 각각 LSTM layer 를 2 개씩 쌓아 총 4 개의 LSTM layer 로 구성하였다.

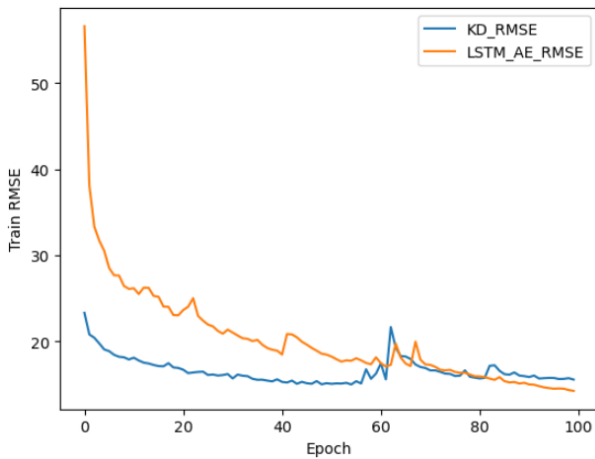
4. 수치 모델을 활용한 지식 증류법 및 결과

본 연구에서는 교사 모델로 수치 모델을 사용하였다. 수치 모델은 기상예보 시뮬레이션을 위해 활용하는 수치 기반 모델을 말한다. 그림 1 은 수치 모델 기반 지식 증류 기법을 보여준다.



[그림 1] 수치 모델 기반 지식 증류법 다이어그램

수치 모델을 활용한 지식 증류법을 적용하였을 때와 적용하지 않았을 때의 성능을 측정하였다. 비교 대상은 이전 연구 [3]에서 미세먼지 예측을 위해 활용한 LSTM 과 AutoEncoder 를 활용한 LSTM_AE 를 대상으로 하였다. 그림 2 에서 보여주듯이, 지식 증류 기법(파란색 선)을 활용하였을 때, 단독 모델 학습(주황색 선)에 비해 훨씬 빨리 동일한 수준의 RSME 에 도달함을 볼 수 있다. 예를 들어, KD_RMSE 의 경우, 20 epoch 에서 이미 20 이하의 RMSE 를 달성하였으며, 이는 LSTM_AE_RMSE 가 80 epoch 에서 달성한 수준과 동일하다.



[그림 2] 기본 모델 LSTM_AE[1]의 Train RMSE(주황색)과 지식 증류를 적용한 모델의 Train RMSE(파란색)

5. 결론

지식 증류법을 적용시킨 모델은 최적의 성능을 더 빠르게 찾는다. 기상 예측 모델에서도 60 이하의 epoch 에선 지식 증류법을 적용시킨 코드가 더 작은 RMSE 값을 가졌다. 기상 예측을 위한 지식 증류는 모델 경량화의 측면에서 이점이 있다.

사사(Acknowledgement)

이 연구는 기상청 “기후 및 기후변화 감시 예측 정보 응용 기술개발 (KMI-2021-01310)” 의 지원으로 수행되었습니다.

참고문헌

- [1] G. Hinton, O. Vinyals, and J. Dean, “Distilling the Knowledge in a Neural Network.” arXiv, Mar. 09, 2015. doi: 10.48550/arXiv.1503.02531.
- [2] K.-K. Ko, Shahzad, and E.-S. Jung, “Big Data Merging and Deep Learning Model Optimization for improving Weather Information Forecasting Performance” Journal of The Institute of Electronics and Information Engineers Vol.58, NO.5, May 2021. doi: 10.5573/iee.2021.58.5.39
- [3] K.-K. Ko and E.-S. Jung, “Improving Air Pollution Prediction System through Multimodal Deep Learning Model Optimization,” Applied Sciences, vol. 12, no. 20, Art. no. 20, Jan. 2022, doi: 10.3390/app122010405.