

LSTM을 이용한 PM_{2.5} 예측 시 발생하는 후행성 문제 연구

김찬수¹, 서명석², 이승희¹

¹국립공주대학교 응용수학과

²국립공주대학교 대기과학과

산업 생태계와 기후 등의 변화로 인해 대기오염의 심각성은 점점 증가하고 있는 상황이며 이들 중에서 미세먼지 농도는 사회 전반에 심각한 문제를 야기하고 있어 정확한 예측을 위해 머신러닝과 딥러닝을 통한 연구가 활발하게 진행되고 있다. 특히 미세먼지 등과 같은 시계열 자료의 경우 딥러닝 모델 중 하나인 LSTM(Long Short-Term Memory) 모델을 사용하여 시계열 예측 연구에 많이 사용되어지고 있다. 그러나, LSTM 모델을 이용하여 미세먼지 등과 같은 시계열 자료를 예측하는데 있어 예측치들이 전일의 값들을 따라 가는 후행성 문제가 발생하고 있다. 본 연구에서는 LSTM 예측 시 나타나고 있는 후행성 문제를 개선하기 위해 디노이징 기법 중 Christiano-Fitzgerald Ban Pass Filter를 통한 노이즈 제거와 이산 웨이블릿 분해를 통해 얻어진 전처리 과정들에 LSTM을 적용하여 후행성 문제를 해결 한 후 두 방법들 간의 예측 성능을 비교 분석하였다. 후행성 개선을 통한 예측 성능을 비교하기 위해 2018년 12월 1일부터 2021년 11월 30일까지 서울시 중구 대기 관측소에서 측정된 시간 당 PM_{2.5} 농도 시계열 자료를 사용하였고, 분석 결과 두 방법 모두에서 후행성 문제가 개선되었음을 알 수 있었고 유사한 예측 성능을 보여주었다. 이러한 후행성 문제 해결을 통해 보다 정확한 예측 성능을 유지하기 위해서는 기온, 시정 등 다양한 시계열 데이터에 적용하여 예측 성능에 대한 비교 분석이 필요한 것으로 여겨진다.

Key words: 노이즈 제거, 미세먼지, 후행성 문제, LSTM(Long Short-Term Memory)