

최적 환경 제어를 위한 LSTM 기반 시간 단위 환경 변화량 예측 시스템

황진주¹, 남태민², 최진철^{3*}, 박찬원³

동의대학교¹, 한국기술교육대학교², 한국전자통신연구원³

{hjj32311¹, ntm32309², spiders22v^{3*}, cwp³}@etri.re.kr

LSTM-based time unit environmental change prediction system for optimal environmental control

Hwang Jinju¹, Nam Taemin², Choi Jinchul^{3*}, Park Chanwon³

Dong-Eui University¹, Korea University of Technology and Education², Electronics and Telecommunications Research Institute³

요약

병원 무균실, 반도체 공장, 양식장 등에서는 미세한 환경 변화에 따라 큰 손실이 발생할 수 있으므로 온/습도, 미세먼지, 조도 등과 같은 환경 변화에 신속하게 대응해야 한다. 이러한 환경 변화를 사전에 예측할 수 있다면, 이상 상태에 따른 피해를 사전에 최소화 할 수 있다. 본 논문에서는 10년 동안의 온도, 습도 기압 실측 데이터를 LSTM 모델로 학습시켜 1시간 단위로 온도/습도/기압을 예측하는 시스템을 제안한다. 제안하는 예측 시스템은 비록 먼 미래에 대한 예측 정확도는 그리 높지 않았으나, 4시간 미만의 예측에서는 비교적 높은 정확도를 보여주었다.

I. 서론

환경 변화에 따른 손실 발생 우려가 있는 공간에서 환경 변화를 사전에 예측할 수 있다면, 피해를 최소화하거나 환경 제어를 위해 사용하는 에너지량을 절감하기 위한 기반 정보로 활용할 수 있다[1]. 병원 무균실, 반도체 공장, 양식장 등에서는 자동화된 IoT 시스템을 이용하여 환경 상태를 관리하고 있다[2]. 그러나 이러한 시스템은 사전에 정의된 간단한 규칙에 의해 제어되거나 사용자의 수동 장치 조작을 통해 환경 상태가 관리되는 경우도 많아 지속적이고 체계적인 환경관리가 어려운 실정이다[3]. 본 논문에서는 환경 변화량을 사전에 예측하기 위해 실측 환경 데이터를 LSTM 모델로 학습시켜 1시간 단위로 온도/습도/기압을 예측하는 시스템을 제안한다.

II. 본론

2-1. LSTM(Long Short Term Memory)

본 논문의 예측 대상인 온도/습도/기압 정보는 시시각각 변하는 시계열 데이터의 특성을 가지기 때문에 LSTM 모델을 통해 학습되었다. LSTM(Long Short Term Memory)이란 순환신경망 구조를 가지며, RNN(Recurrent Neural Network)의 장기 의존성 문제를 개선한 모델이다[4]. 단위시간 별 온도, 습도, 기압이 모델을 입력변수로 각각 이용되고 1~12 시간 이후의 예측 온도, 습도, 기압이 그 출력변수로 도출된다.

2-2. 학습 데이터

기상청에서 제공하는 2012년부터 2022년 8월까지의 1시간 단위의 양산시의 온도/습도/기압 데이터 88,382개를 선행 학습 모델 생성에 이용하였다[5]. 그림 1과 같은 시계열 데이터를 0.85 : 0.15 비율로 나누어 각 75252/13280 개의 데이터를 이용해 학습과 검증을 진행했다. 이후 오픈 API를 이용해 현행화된 데이터를 수집해 학습 모델을 업데이트했다.

2-3. 시스템 구성

사용자가 웹 GUI를 통해 희망하는 예측 시간을 입력하면, 시스템은

LSTM 학습 모델을 통해 추론하고, 추론된 결과를 가시화한다. 시스템의 추론 동작 시퀀스는 다음 그림 1과 같다.

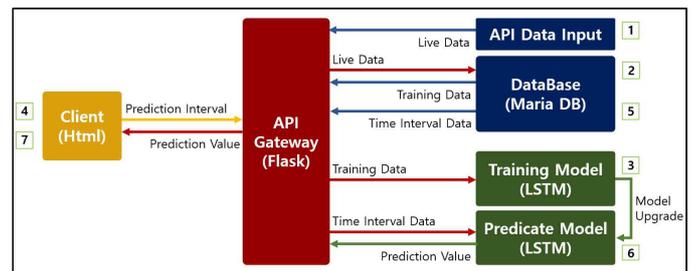


그림 1. 시스템 동작 시퀀스

- ① 오픈 API를 통해 실시간 온도/습도/기압 데이터를 수집한다.
- ② 수집된 데이터를 DB에 저장한다.
- ③ 신규 수집 데이터를 이용하여 기존 LSTM 모델을 추가 학습한다.
- ④ 사용자가 현재로부터 예측하고자 하는 시간 간격을 입력한다.
- ⑤ 입력받은 시간 간격으로 DB에서 시계열 데이터를 추출한다.
- ⑥ 전달받은 시계열 데이터를 LSTM 모델에 입력해 예측값을 도출한다.
- ⑦ 최종 예측된 값을 웹 GUI에 출력한다.

2-3. LSTM 기반 예측 모델

본 연구에서는 예측 시점부터 1시간 간격으로 최대 12시간까지의 온도/습도/기압을 예측할 수 있는 12개의 학습 모델을 도출했다. 각 예측 모델은 과거 200개의 데이터 샘플을 참조해 예측이 가능하도록 설정했다. 예측 모델이 참조할 데이터는 그림 2와 같이 세 가지 형태로 가정하였다.

첫 번째 가정은 (1년 전 시각+희망 예측 시각)에서 과거 200개의 연속 데이터를 참조하는 것이다. 예를 들어 22년 8월 13일 16시의 2시간 후의 예측 결과를 확인하고 싶다면, 21년 8월 13일 18시를 기준으로 과거 200개의 연속 데이터를 참조하여 예측하는 것이다. 두 번째 가정은 현시점에서 과거 200개의 데이터를 참조하는 것이다. 마지막으로 세 번째 가정은 과거 200개의 이산 데이터를 참조하는 것이며, 각 이산 데이터의 간격은 예

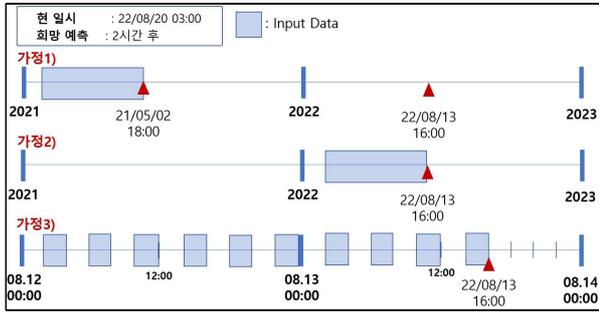


그림 2. 예측을 위한 입력 date 구상

측하고자 하는 시간과 동일하다. 예를 들어 8월 13일 16시에 2시간 후의 미래를 예측하고 싶다면, 2시간 간격인 16, 14, 12, 10의 시간 간격의 데이터를 참조한다.

III. 성능 분석

본 시스템에서는 가장 좋은 성능을 보인 가정 3의 이산 데이터를 참조하여 성능을 분석하였다. 표 1에 예측 모델별 온도/습도/기압 예측 오차를 나타내었다. 표 1 상에 오차는 Mean Absolute Error로써 실제값과 예측값에 대한 평균 절대 오차값이다. 본 성능시험에서는 각 모델별로 1, 2, 4, 8시간 후를 예측한 오차를 도출했다. 예를 들어 1시간 후 예측 모델에 희망 예측 시간이 2시간이라는 것의 의미는, 1시간 후의 온도/습도/기압을 예측할 수 있도록 학습시킨 모델을 이용해 2시간 후를 예측하는 것이다. 1시간 후 예측 모델의 경우 1, 2, 4, 8시간 후의 온도를 예측한 오차는 각각 0.4789, 1.5651, 3.3006, 4.1446로써, 예측 시간이 증가할수록 오차도 함께 증가하였다. 마찬가지로 2, 4, 8시간을 예측하는 모델들도 전반적으로 예측 시간이 증가할수록 오차도 함께 증가하는 특성을 보였다.

반면에, 희망 예측 시간을 기준으로 하는 희망 예측 시간과 n시간 후 예측 모델이 동일한 경우 오차가 가장 낮았다. 1시간 예측을 한 경우에는 1시간 후 예측 모델을 사용했을 때 오차가 가장 낮았으며, 2, 4, 8시간을 예측한 경우에도 2, 4, 8시간 모델을 사용한 경우가 오차가 가장 낮았다. 따라서 각 예측 모델은 동일한 희망 예측 시간으로 활용할 때 가장 높은 정확도로 사용될 수 있다.

현재 모델에서는 시간 간격이 커질수록 steps_per_epoch과 epoch을 증

표 1. 희망 예측 시간 별 참조 데이터에 따른 온도 예측 오차

n시간 후 예측 모델	희망 예측 시간	오차		
		온도	습도	기압
1 hour	1 hour	0.4789	2.8653	0.4445
	2 hour	1.5651	5.4450	0.4014
	4 hour	3.3006	12.4648	1.3702
	8 hour	4.1446	21.5764	1.6678
2 hour	1 hour	0.9037	4.1899	0.4577
	2 hour	0.9751	4.6113	0.3892
	4 hour	2.9298	10.9491	1.1498
	8 hour	3.5407	20.8980	2.2336
4 hour	1 hour	0.6797	6.0997	0.4417
	2 hour	1.1199	4.0003	0.4985
	4 hour	2.3121	8.3772	0.9227
	8 hour	3.3457	15.5100	1.4559
8 hour	1 hour	0.6353	4.4487	0.5809
	2 hour	1.1996	4.8721	0.7156
	4 hour	3.4174	14.3692	1.0508
	8 hour	2.6069	14.6309	1.3623

가해 더 오랜 학습을 하도록 하였다. 1시간 간격의 모델에서는 2000/40으로 두고 1시간마다 각각 100/10의 값을 더하여 작동하였다. 하지만 그림 3을 통해 알 수 있듯이 많은 학습을 하여도 완전한 극복이 되지 않음을 알 수 있다.

2-4. 처리 정확성

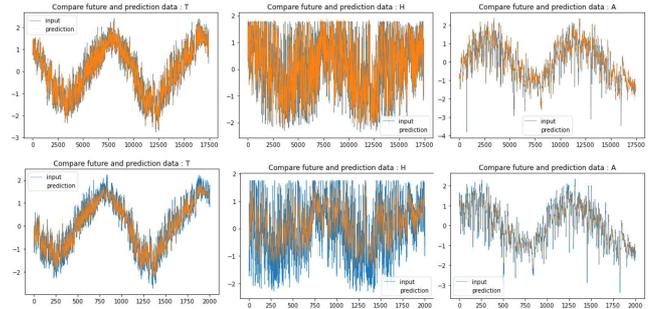


그림 3. 예측 모델에 따른 예측 데이터와 실제 데이터

그림 3은 예측 모델에 따른 예측 데이터와 실제 데이터를 비교한 그래프이다. 위의 세 그래프는 1시간 간격의 데이터를 학습한 모델의 예측 결과며, 아래 3개의 그래프는 8시간 간격의 데이터를 학습한 경우를 보인다. 1시간 간격의 모델은 온도를 기준으로 최대 오차 값이 0.8에 미치지지만 8시간 모델은 1에서 6까지의 값을 보여 낮은 정확도를 보여주었다.

IV. 결론

본 논문에서는 온도/습도/조도의 단일 속성을 LSTM 모델에 학습해 1시간 단위 이후의 온도/습도/조도를 예측하였다. 4시간 이후의 예측에서는 비교적 낮은 정확도를 보였지만, 이전 1시간까지의 예측에서는 높은 정확도를 보였다. 향후 강수, 풍속 등 연관되는 데이터를 추가하여 multivariate 시계열 데이터를 이용한 추론법을 적용하여 예측 결과에 대한 정확성을 높일 것이다.

ACKNOWLEDGMENT

이 논문은 2022 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 100%지원을 받아 수행된 연구임. (No. 2022-0-00438, (총괄,1 세부) 지능형 디지털 연합트윈 운용 및 예측 핵심기술 개발)

참고 문헌

- [1] 최중석, 신용태, “매장 에너지 절감을 위한 LSTM 기반의 전력부하 예측 시스템 설계,” 한국정보전자통신기술학회 논문지 제14권 제4호, pp. 307 - 313, 2021.
- [2] Nuvvula, Adiraju, Mubin, Valisetty and B. Shahana, “Environmental smart agriculture monitoring system using internet of things,” International Journal of Pure and Applied Mathematics. Vol. 115. pp. 313-320, 2017.
- [3] 이정현, 김형욱, 김동희, “사용자가 원격 제어하는 블루투스 기반 스마트 홈 IoT 시스템 구현,” 디지털콘텐츠학회논문지 제22권 제2호, pp. 331-338, 2021.
- [4] 주일택, 최승호, “양방향 LSTM 순환신경망 기반 추가예측모델,” 한국정보전자통신기술학회논문지, 제11권 제2호, pp. 204-208, 2018
- [5] 기상청 기상자료개방포털, 종관기상관측(ASOS), URI: <https://data.kma.go.kr/data/grnd/selectAsosRltmList.do?pgmNo=36>, 2022.