

효율적인 작물 재배를 위한 LSTM 기반 작물 잎 중량 예측 모델

이민서, 장혜원, 이예람, 김현희
동덕여자대학교 정보통계학과

20191059@dongduk.ac.kr, 20201071@dongduk.ac.kr,
20201066@dongduk.ac.kr, heekim@dongduk.ac.kr

LSTM-based crop leaf weight prediction model for efficient crop cultivation

Lee Min Seo, Chang Hye Won, Lee Ye Ram, Kim Hyon Hee
Dept. of Information Statistics, Dongduk Women's University

요 약

지구온난화와 농업 활동 감소로 인해 농작물 생산량이 줄어드는 추세이다. 4차 산업혁명 시대를 맞아 농업 분야에서도 인공지능 기술을 활용하여 효율적인 작물 재배가 가능해지고 있다. 작물의 수확량을 최고로 끌어올릴 수 있는 시간대별 최적 환경을 알아낼 수 있다면 식물 재배와 관련한 제반 사업에 도움이 될 것이다. 본 연구에서는 LSTM 알고리즘을 이용하여 상추의 일별 중량을 예측하는 인공지능 모델을 생성하였다. 제안하는 AI 예측 모델을 통해, 보다 효율적인 작물 재배가 가능해질 수 있을 것으로 보인다.

1. 서론

지구온난화로 인해 홍수와 가뭄이 공존하는 물 공급 불균형은 농작물 생산에 지장을 초래한다. 또한 농촌의 농가 인구, 농가 수, 토지 면적 감소가 원인이 되어 농작물 생산량이 감소하고 있다.[1] 이에 효율적인 작물 재배의 필요성이 대두되고 있어 농업 분야에서도 인공지능 기술이 사용되고 있는 추세이다.

본 연구는 작물의 수확량을 최고로 끌어올릴 수 있는 시간대별 최적 환경을 알아내어, 작물의 일별 잎 중량을 예측하는 AI 예측 모델을 생성하는 것에 목적을 둔다. 제공받은 시계열 데이터를 바탕으로 필요한 변수들만 선정한 뒤, 정규화를 사용하여 변수들의 단위를 표준화 시켜준다. 데이터는 시간 단위의 한 달 치 input 데이터와 날짜 단위의 한 달 치 target 데이터로, target 데이터를 input 데이터와 동일한 시간 단위의 데이터로 만들어 주기 위해 선형 보간법을 적용한다. 전처리된 데이터를 바탕으로 LSTM을 적용하여 모델 학습을 진행했고, 가장 좋은 성능을 지닌 모델을 생성했다.

생성된 모델을 통해 작물의 일별 최대 잎 중량을 알 수 있게 되었고, 이를 바탕으로

최적의 생육 환경도 조성함으로써 농작물 생산량 증가에 도움을 줄 수 있을 것으로 보인다.

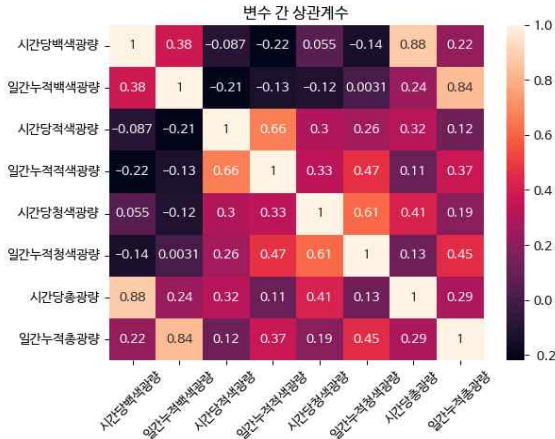
본 논문은 다음과 같이 구성된다. 제2장에서는 데이터 수집 및 전처리를 나타내고, 제3장에서는 모델 생성 및 분석 결과를 나타낸다. 마지막으로 제4장에서 결론을 맺는다.

2. 데이터 수집 및 전처리

본 연구는 2020년도 정부(과학기술정보통신부) 재원으로 한국연구재단의 지원을 받아 『한국과학기술연구원(KIST)-코리아디지털(주)-주식회사 파모스』에서 수행한 『개방형 식물 표현체 데이터 수집-공유-분석-활용 플랫폼 개발』 연구를 통해 도출한 데이터로 진행하였음을 명시한다. 28개의 상추 관측치가 있으며 모델에 사용될 input과 예측해야 할 target으로 구성된다. input은 (생육일, 관측 시간) 별로 내부 온도, 내부 습도, CO₂, EC, 시간당 분무량, 일간 누적 분무량 등 총 14개 변수이며, target은 생육일과 일별 잎 중량으로 이루어져 있다.

한 개의 상추 케이스당 input은 시간별 데이터로 28(일)*24(시간)의 672개 데이터이지만, target은 일별 데이터로 28개의 데이터이기 때문에 input과 target의 데이터 구조가 맞지 않다.

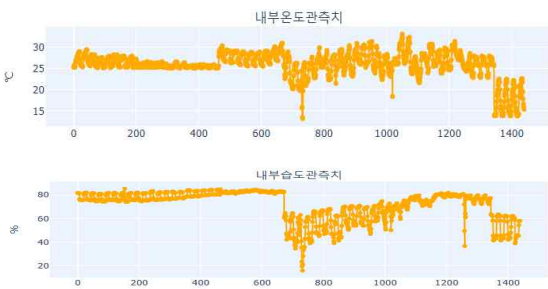
target에 맞추어 시간별 데이터를 일별 데이터로 변형한다면 28개의 데이터 손실과 시계열의 특성 미반영 문제로 input 데이터 형식에 맞추어 진행한다. 따라서 선형보간법[2]을 이용해 target 데이터를 시간별 데이터로 바꾸고, 변수들의 단위를 맞추기 위해 정규화를 진행한다.



(그림 1) 광량 변수들 간의 상관계수

작물의 수확량을 최대로 하는 시간대를 찾는 것이 목적이기 때문에 (그림 1)을 통해 무의미한 광량 값이 많다는 것을 알 수 있어 적색, 백색, 청색을 합한 시간당 총광량만을 사용한다. 따라서 최종 사용된 변수는 내부 온도, 내부 습도, co2, EC, 시간당 총광량 5개이다. 관측된 기간이 28일로 짧기에, 28개의 상추 관측치를 하나로 이어 붙인다.

3. 모델 생성 및 분석 결과

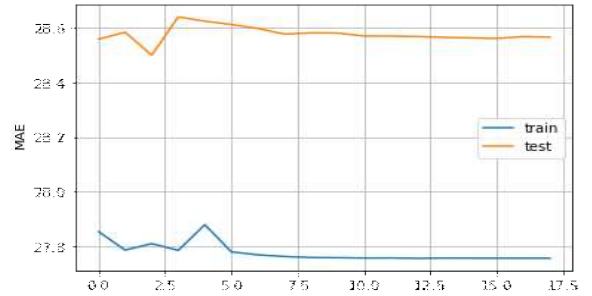


(그림 2) 주기를 가지는 변수들

(그림 2)와 같이 상추 관측치 별로 주기를 가지고, 한 상추 개체 안에서도 시간별로 측정하였기 때문에 시계열의 특성을 가진다. LSTM 알고리즘[3]으로 상추의 일별 잎 중량을 예측했다.

여러 가지 활성화 함수를 적용해 본 결과, 가장 좋은 성능을 보인 tanh 함수를 택했다. 모델의 가중치를 제한해서 일반화 성능을 향상시키는

데 사용되는 L2 정규화를 사용하여, Fully-Connected 층에 LSTM의 출력 벡터를 입력한 모델을 완성함으로써, LSTM 층의 정확한 출력 예측이 가능했다. 손실 함수, 평가 지표는 mae, 최적화 알고리즘은 adam을 사용하여 이상치에 영향을 덜 받도록 구성했다.



(그림 3) 최종 선정 모델 성능 시각화

따라서 활성화 함수로 tanh, 과적합 방지와 일반화 성능 향상을 위한 dropout 기법, L1 정규화를 사용하여 예측한 성능은 (그림 3)과 같이 loss 값 28.57, mae 값 28.57이 나왔다.

4. 결론

본 연구에서는 상추의 일별 잎 중량을 예측하는 AI 예측 모델을 생성하여 작물의 수확량을 최대로 끌어올릴 수 있는 시간대별 최적 환경을 예측했다. 짧은 관측 기간과 중복되는 변수들로 인해 예측 오차가 높은 편이라, 기상 데이터와 작물에 관한 연구를 토대로 정확한 연구를 모색 중이다. 최종 모델을 기반으로 한 예측 모델을 이용해서 생성 모델을 만든 후, 상추의 일별 잎 중량 예측값을 기반으로 하여 상추가 자랄 수 있는 최적의 생육 환경 조성 수치를 예측할 예정이다.

다양한 원인으로 인해 농작물 생산량이 크게 감소한 현시점에, 연구 결과인 AI 예측 모델을 기반으로 추후 생성 모델도 만든다면 효율적인 작물 재배가 가능해져 농작물 생산량을 증가시킬 수 있을 것으로 기대한다.

참고문헌

[1] 이기섭, “기후변화가 농작물 재배환경에 미치는 영향”, 광주 : 전남대학교 산업대학원, 2014, 17-21
 [2] 최인하, 차경환, 김경민, 안중화, “선형보간법을 이용한 기계학습 기반 매립지 침출수 발생량 예측 모델 개발”, 대한환경공학회지, 2023, 1, 11-20, 45.
 [3] Sagarika Sharma, Sujit Rai, Narayanan C. Krishnan, “Wheat Crop Yield Prediction Using Deep LSTM Model”, 2020, 1-7