

기계 학습을 이용한 금강유역 옥천의 오염부하량 예측 Prediction of pollution loads in Geum River using machine learning

임희성*, 안현욱**
Heesung Lim, Hyunuk An

요 지

기후변화에 따른 환경오염은 21세기 인류에게 가장 심각한 문제 중의 하나로 대두되고 있다. 환경적인 측면에서 하천오염은 경제적으로 많은 문제를 발생시키고 있다. 이러한 하천오염 문제를 해결하기 위해서는 오염물질의 농도 측정 및 데이터 측적이 필수적이라 할 수 있다. 그러나 일반적으로 오염물질 부하량에 대한 직접적인 측정은 비용 측면에서 쉽지 않은 것이 사실이다. 또한 실시간으로 BOD, COD, TN, TP 등의 자료를 이용하여 예측하는 것에는 자료의 부족성으로 인해 한계가 있다.

본 연구에서는 구글의 딥러닝 오픈소스 라이브러리인 텐서플로우를 활용하여 기계 학습을 통한 하천오염 예측을 목적으로 하고 있다. 기계 학습을 위하여 텐서플로우를 활용하여 RNN, LSTM 인공신경망 모형을 구축하였다. 하천오염의 학습과 예측을 위해 결과치 분석을 위한 자료로는 금강유역에 위치한 옥천 관측소 충청북도 옥천군 이원면 이원대교에 위치한 36°14'31.0"N 127°40'02.6"E의 관측소에서 BOD, COD, DO, 부유물질의 자료를 사용하였다. 모형의 학습을 위해서 입력자료는 수위, 유량, 평균기온, 평균풍속 자료를 2004년 ~ 2017년까지의 14년간의 자료를 사용하였다. 연구를 위해 BOD, COD, DO 부유물질 자료는 물환경정보시스템 (<http://water.nier.go.kr/>)의 자료를 활용하고 수위, 유량등의 자료는 국가수자원관리종합정보시스템 (<http://www.wamis.go.kr/>)의 자료를 사용하였다. 그러나 수온, 수위, 풍속등의 자료는 일 자료가 있는가 반면 BOD, COD, TN, TP등의 자료는 일 자료가 있지 않아 이를 원활히 활용할 수 있도록 예측을 위한 결과치의 선형보간법을 통해 일 자료를 획득한 후 연구를 하였다. RNN, LSTM의 분석 시 학습속도, 반복시행횟수 sequence length의 길이 등의 값을 조절 하면서 결과치를 분석하였다.

핵심용어 : 기계 학습, RNN, LSTM, 오염부하량

* 정회원 · 충남대학교 농공학과 수리환경 및 자원정보 석사과정 임희성 · E-mail : dlagmltjd123@naver.com

** 정회원 · 충남대학교 지역환경토목학과 교수 안현욱 · E-mail : hyunuk@cnu.ac.kr