

## RNN 알고리즘을 이용한 온천천의 하천수질 인자 예측

Prediction of river water quality factor at Oncheoncheon Basin using RNN algorithm

임희성\*, 안현욱\*\*

Heesung Lim, Hyunuk An

### 요 지

인구의 도시 집중화로 인하여 다량의 생활용수의 사용에 따라 하천의 자정능력을 초과하여 오염을 유발시키고 있다. 이에 도시하천들의 오염은 점점 심해져 경제적으로 많은 문제를 유발하고 있다. 이러한 하천오염 문제를 과학적으로 대응하기 위해서는 오염물질의 농도 측정 및 데이터 축적을 통한 오염예측이 필수적이라 할 수 있으며, 부산광역시 보건환경정보 공개시스템에서는 하천수질 자동측정망을 설치하여 시간 단위로 오염물질을 측정하고 있다. 그러나 온천천의 하천수질 데이터는 계속 쌓여가고 있는데 이 데이터를 활용해서 하천수질 인자 예측이 거의 이뤄지지 않고 있다.

본 연구에서는 순환신경망 알고리즘을 활용하여 일 단위의 하천수질 인자 예측을 시도하였다. 순환신경망은 인공신경망의 발전된 형태인 시계열 학습에 강한 RNN, LSTM 알고리즘을 활용한 일 단위 하천수질 인자 예측을 하고자 하였다. 연구에 앞서 시간 단위로 쌓여있는 데이터를 평균 내어 일 단위로 변경하였고 이 데이터를 가지고 일 단위 하천수질 인자 예측을 진행하였다. 연구에는 Google에서 개발한 딥러닝 오픈소스 라이브러리인 텐서플로우를 활용하여 DO, 탁도 등 항목을 예측하였다. 하천오염의 학습과 예측을 위해 대상지로는 부산지역 온천천의 부곡교, 세병교, 이섭교 관측소를 선택하였다. 연구를 위해 DO, 탁도 등 자료 수집은 부산광역시 보건환경정보 공개시스템의 자료를 활용하였다. 모형의 학습을 위해 입력자료로는 하천수질 인자 자료를 이용하였고, 자료의 학습에는 2014년~2017년 4년간의 자료를 학습자료로 사용하였고, 2018년 1년간의 자료는 모형의 검증에 사용하였다. RNN, LSTM 알고리즘을 활용하여 분석 시 은닉층의 개수, 반복시행횟수, sequence length 등의 값을 조절하여 하천수질 인자 예측을 하였다. 모형의 검증을 위해  $R^2$ (r square)와 RMSE(root mean square error)를 이용하여 통계분석을 실시하였다.

**핵심용어 : 하천수질, RNN, LSTM, 기계학습**

\* 정회원 · 충남대학교 지역환경토목학과 수리환경 및 자원정보 박사과정 E-mail : hslim1@cnu.ac.kr

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