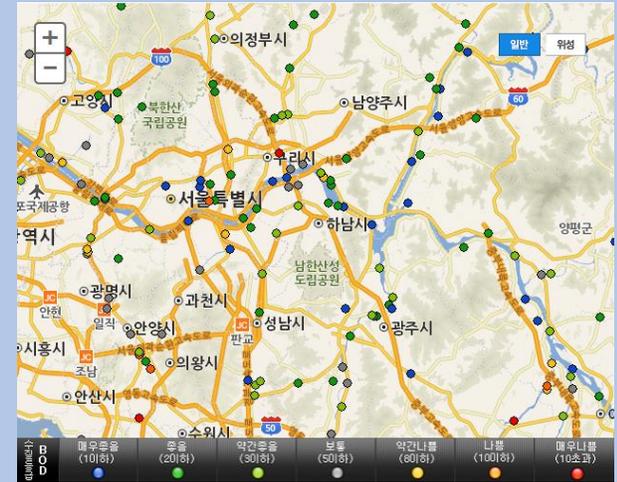


데이터 기반 한강 수질 예측

홍한움

연구목적 및 필요성

- 측정 전 수질을 예측을 통한 효율적인 수질 관리
- 이공계 모형 (QUAL2K, WASP, etc)의 한계
 - 하천 및 호수를 소구역으로 분할하고 각 소구역에 대한 오염물질 발생부하량을 자료에 투입해야 하는데, 공간 해상도를 만족하는 방대한 자료를 구하기 어려움
- 최근 인공지능 기법과 공간통계모형의 발달로 데이터기반 수질 예측 가능
- 특정 하천에 국한되지 않은 예측모형 적용 가능



선행연구

- 한국정보화진흥원
 - 낙동강 유역 녹조 발생 예측
 - SVM, Random forest, RNN, 선형회귀모형
- Rankovic et al. (2012)
 - 딥러닝 기반 Gruža 저수지 용존산소량 예측
- Xu et al. (2012)
 - 시공간분석 기반 Zhangweinan 강 용존산소량 분석

Datasets

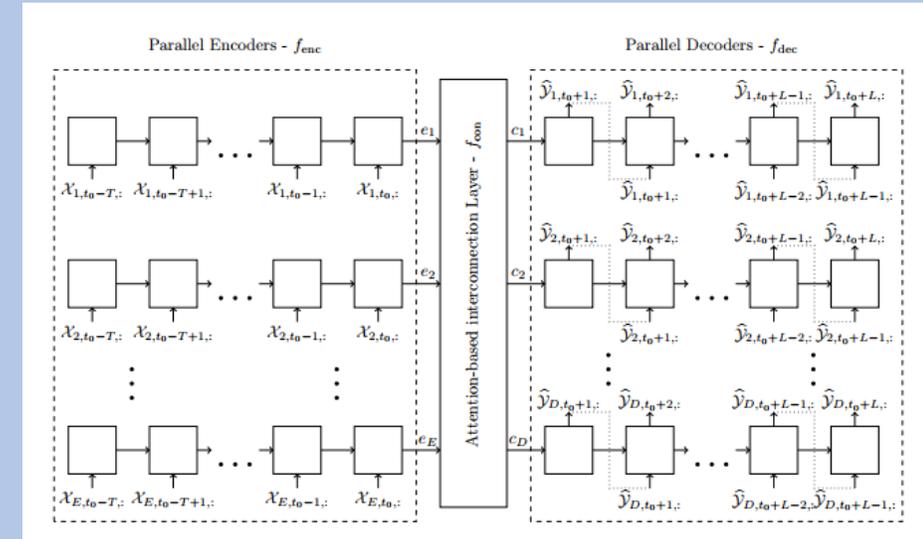
- 수질 일반측정망 (from 물환경정보시스템), 기상자료
 - 수소이온농도(pH), **용존산소량(DO)**, BOD, COD, 부유물질, 총질소(TN), 총인(TP), TOC, 수온, 전기전도도, 총대장균군수, 용존총질소, **암모니아성질소(NH3-N)**, 질산성질소(NO3-N), 용존총인, 인산염인, **클로로필-a**, 분원성대장균군수, **강우량**, 습도, 기온
- 오염원자료의 활용가능성
 - 현재 연해상도의 자료만 있어 활용 불가
 - 위치정보와 함께 주 혹은 월 해상도로 자료가 있을 경우 주요 설명변수로 활용 가능함
- 2010-2017 수도권 수질 측정지역 주 해상도 자료이용

가용자료 개수	측정소 수
400이상	3
350-400	9
300-350	24
200-300	21
151미만	159

- 350개 이상의 자료를 사용 가능한 12개 측정소 대상으로 예측 진행
- 나머지 자료는 설명변수로 활용

Methodology

1. ANN
2. RNN, GRU (or LSTM)
3. KNN
4. 시공간자료분석모형
5. AutoEncoder



- 분석언어: python tensorflow
- 선형회귀모형, VARMA 대비 우수한 예측 결과 예상
 - 자료의 비선형성, 공간정보활용성

예상소요기간

- 자료수집 및 전처리 ; 3월
 - 도메인 분석
 - 결측값 보간
 - 이상치 제거/대체
- 자료분석 ; 4-8월
 - Simple ANN, RNN ; 4-6월
 - GRU, KNN ; 5-6월
 - 시공간자료분석 ; 5-6월 //중간보고
 - AutoEncoder ; 7-8월
- 보고서 작성 ; 6-10월