

06

chapter

AI 기반 빅 데이터 분석 적용 사례 〈IoT기반 환경보건〉

조재혁 || 송실대학교 교수, AI데이터 연구단장
 금윤재 || 송실대학교 선임연구원

I. 개요

인공지능(AI)은 여러 학문이 연계된 융합 분야이며 기술적 관점에서는 인간의 인지, 학습, 추론 등 지적 능력을 기계(컴퓨터)로 구현하는 기술로 정의한다.[1] 인공지능은 데이터와 지식이 산업 경쟁력의 핵심 원천이 되는 4차 산업혁명의 핵심 원동력이며 의심할 여지없이 현재 가장 주요한 기술 트렌드 중 하나이다. 따라서 세계는 AI 기술 수준이 국가의 역량을 좌우하는 핵심 기술로 보고 있다. 특히 머신러닝 알고리즘은 우리가 원하는 부분의 솔루션에서 더욱 정교해지고 정확해지고 있다. 이는 공공 영역의 데이터를 활용하여 컴퓨터 비전 시스템에서부터 접촉 추적 이니셔티브를 통해 밝혀진 상호작용을 분석하는 데 이르기까지 자기 학습 알고리즘은 수동적인 인간 분석을 통해 눈에 띄지 않는 연결과 통찰력을 발견할 것이다.[2]

세계적인 IT 시장조사기관인 가트너는 COVID-19 위기에 대응하고 성장 추세로 변화하기 위한 전략에 주목했으며 중대한 격변과 세계의 경제현황 속에서 탄력적인 사업체들이 필요로 하는 9대 전략적 기술 트렌드를 발표하였다. [그림1]에서 보면 2021년 AI 전략기술 트렌드는 사람중심으로서 사람과 기술 간의 상호작용에 초점을 맞추고, 여러 기술이 통합을 강조하였다.[3] 또한 인공지능과 함께 새로운 비즈니스 모델인 AI 엔지니어링, 초 자동화를 만들어 낼 것이라는 전망을 내 놓았다.



〈자료〉 “I-Korea 4.0 실현을 위한 인공지능(AI) R&D”과학기술정보통신부, 2018. 5

그림 2. 인공지능 기술 개요도

[그림 2]에서 보면 AI의 도입은 안전, 의료, 환경(농업, 도시, 에너지) 등 다양한 분야에 적용되고 있음을 보여준다. 특히 환경오염(미세먼지, 유해물질 등)문제와 국제적인 팬데믹상황으로 AI의 활용이 보편화될 가능성이 클 것으로 전망했다. 정보통신정책연구원은 최근 발표한 보고서[6]에 따르면 보건 의료분야에서는 AI 성과가 상당히 높고 상당수준의 기술을 선보였다고 소개했다. 국내에서는 AI을 기반으로 성남시는 네이버 클로버 AI 음성봇으로 감염 의심자 감시, 전주시는 한컴 AI체크 25를 이용하여 자가격리자 관리, KT는 코로나 확산 예측 연구에 AI를 사용하였다.

2016년 다보스포럼에서 발표한 상위 ‘10대 글로벌 리스크 중 단 한번의 실패로도 가장 큰 영향을 줄 수 있는 것’은 환경을 꼽았다. 또한 마이크로소프트는 지구환경 분야에 인공지능을 사용하는 ‘지구환경AI(AI for Earth)’ 프로젝트를 출범하였다. 또한 과학기술정보통신부와 국토교통부에서 2017년도부터 ‘스마트 환경 모니터링 실증사업’을 추진하여 실시간으로 발생하는 악취, 미세먼지 등을 통해 신속한 원인 규명 및 대응을 통해 환경 사각지대의 해소에 기여하고 있다. 이처럼 환경분야는 빅데이터와 AI의 활약이 가장 절실한 분야이다.[7]

본 고에서는 AI 기반 빅데이터 분석의 개발 동향과 국내외의 IoT기반 환경보건분야의 기술 연구 및 사례와 향후 전망에 대해 살펴보고자 한다. 이를 위해 먼저 II 장에

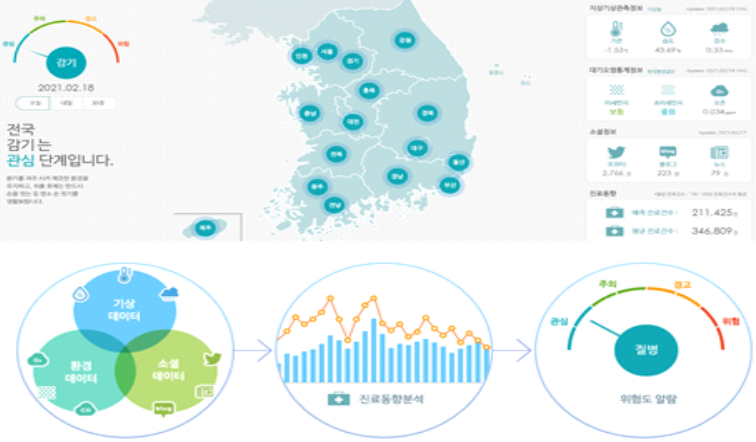
서는 IoT 기반 환경보건 기술 동향과 선행연구들을 보여준다. III 장에서는 IoT기반의 환경보건분야에서 실제 적용되고 있는 활용 사례를 소개함으로써 AI 도입에 따른 기대효과를 확인해 보고, IV장에서 본 고의 결론 및 시사점을 제시한다.

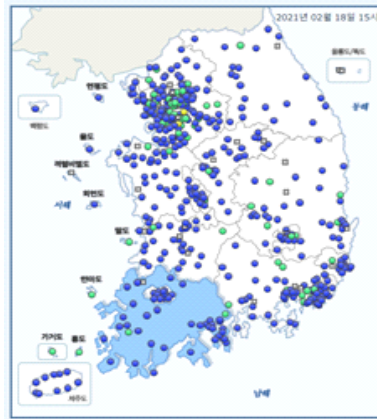
II. IoT 기반 국내 및 국외 환경보건분야 적용사례

1. 국내

[표 1]은 국내 IoT 기반 환경보건분야의 적용 사례를 보여준다. 빅데이터 관련하여 미세먼지, 공기질의 환경 난재의 예측 및 분석이 가능해야하며 다분야에서 빅데이터 분석 R&D를 추진하여야한다. 이는 IoT 정보를 활용하여 실시간으로 얻어진 환경 데이터를 수집 후 데이터 전처리 및 분석할 수 있는 다양한 기술을 확보해야한다.

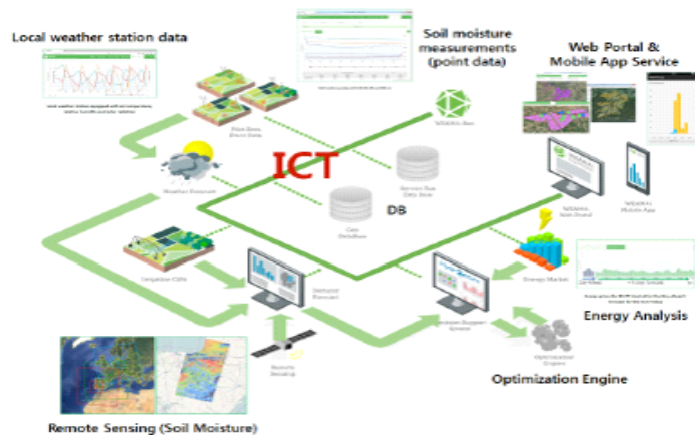
표 1. 국내 IoT 기반 환경보건 분야 적용 사례

구분	주요 내용
국민건강보험공단 <국민건강 알람서비스>	 <p>타 기관이 보유한 식중독, 기상기후, 환경자료를 연계하고 인간의 SNS 정보를 융합하여 주요 질병의 위험도 동향과 알람 제공 서비스</p>
한국환경공단 <실시간 대기오염도 모니터링>	<p>대기질 모니터링 시스템으로 대기환경기준 물질 6개 항목(아황산가스, CO, NO2, 미세먼지 (PM10, PM2.5), 오존)에 대한 대기오염도를 모니터링 하는 시스템. 이는 인체 영향과 체감오염도를 반영한 통합대기환경지수 적용, 4개 등급으로 표현</p>



지능정보산업협회
<저에너지, 저탄소
지향의 지능형 빌딩
관리 플랫폼>

실시간으로 건물 내 에너지 소비현황을 모니터링 함으로써 저에너지, 저탄소를 구현할 수 있는 지능형 빌딩 건설이 확산관리하는 플랫폼이다. 이는 건물 내 환경 및 전력에너지 관련 데이터를 수집하는 에너지 IoT 정보와, 효율적으로 관리하기 위한 에너지 빅 데이터 분석정보를 사용자의 스마트폰에 증강현실(AR)로 구현할 수 있음



<자료 >ETRI - 직독식 수질복합센서 및 초분광영상 기반 시공간 복합 인공지능 녹조 예측 기술 보고서 2019. 5

2. 국외

[표 2]는 국외 IoT 기반 환경보건분야의 기술 및 정책현황 적용 사례를 보여준다. 대기, 수질 등의 다양한 환경 오염도를 측정할 수 있는 GIS 기반 시스템을 도입하였으며 다양한 환경오염 데이터베이스를 수집하고 있다.

표 2. 국외 IoT 기반 환경보전 분야 적용 사례

구분	주요 내용						
<p>미국 <환경보호청 (EnviroAtlas)></p>	<p>주정부, 지역단체, 그리고 민간단체 등에서 정책 또는 개발사업에 대한 영향을 평가할 수 있도록 만들어진 인터넷 기반 환경보전 정보시스템</p>  <p>1) 2014년 미국 환경보호청이 상용화한 환경보전관련 빅 데이터 GIS 도구 2) 대기, 수질, 토양과 같은 기본정보와 자연재해빈도, 재해예방시설, 기후변화, 여행정보와 같은 다양한 정보 제공 3) 현재 환경오염정도에 따른 위해도와 질병발생 가능성 모의 가능</p>						
<p>미국 <환경보건 감시 프로그램></p>	<p>환경 유해성, 건강영향에 대한 자료를 감시 네트워크를 통해, 수집, 통합하여 모든 이해관계자들에게 공개함으로써 유해인자로부터 건강피해를 예방하고 공중 보건을 개선하는 체계 구축 시스템</p>  <div data-bbox="764 1038 1253 1350"> <h3 style="text-align: center;">환경보건 감시 프로그램의 대상</h3> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 33%;">건강 영향</th> <th style="width: 33%;">인구집단 건강</th> <th style="width: 33%;">환경</th> </tr> </thead> <tbody> <tr> <td> <ul style="list-style-type: none"> • 천식 • 선천적 기형 • 열 스트레스 • 암 • 심장질환 </td> <td> <ul style="list-style-type: none"> • 인구집단 특성 (인구 조사) • 생활방식 위해 요소 • 생체모니터링 </td> <td> <ul style="list-style-type: none"> • 기후변화 • 지역사회 설계 • 주거 환경조건 • 실외 공기질 • 지역 수질 </td> </tr> </tbody> </table> </div>	건강 영향	인구집단 건강	환경	<ul style="list-style-type: none"> • 천식 • 선천적 기형 • 열 스트레스 • 암 • 심장질환 	<ul style="list-style-type: none"> • 인구집단 특성 (인구 조사) • 생활방식 위해 요소 • 생체모니터링 	<ul style="list-style-type: none"> • 기후변화 • 지역사회 설계 • 주거 환경조건 • 실외 공기질 • 지역 수질
건강 영향	인구집단 건강	환경					
<ul style="list-style-type: none"> • 천식 • 선천적 기형 • 열 스트레스 • 암 • 심장질환 	<ul style="list-style-type: none"> • 인구집단 특성 (인구 조사) • 생활방식 위해 요소 • 생체모니터링 	<ul style="list-style-type: none"> • 기후변화 • 지역사회 설계 • 주거 환경조건 • 실외 공기질 • 지역 수질 					
<p>EU(포르투갈, 스페인, 독일) <WEAM4i(Water & Energy Advanced Management For Irrigation) 프로젝트></p>	 <p>The diagram illustrates the WEAM4i project architecture. It shows a central 'ICT' hub connected to various data sources and services. On the left, 'Local weather station data' and 'Remote Sensing (Soil Moisture)' feed into the system. On the right, 'Soil moisture measurements (point data)' and 'Energy Analysis' are processed. The 'Optimization Engine' is at the bottom, which then feeds into 'Web Portal & Mobile App Service' at the top right. A 'DB' (Database) is also shown in the center, connected to the ICT hub.</p>						

3개국에서 시범적으로 시행한 프로젝트(2013.11~2017.04)로 관개의 수자원 관리에서 수자원 에너지 측면을 고려하는 혁신적인 솔루션의 개발 및 배포하는 프로젝트. 위성영상의 토양수분 정보, 토양수분 실측값 및 기상자료에 기초한 관개 정보 등을 데이터화하고 이 정보를 이용하여 가뭄 현황을 분석, 살수 관개 장치인 피벗 구동을 위한 전원 공급 정보 등의 솔루션을 개발

영국
<환경 보건 감시 시스템>

아프리카
<AfSIS 프로젝트>

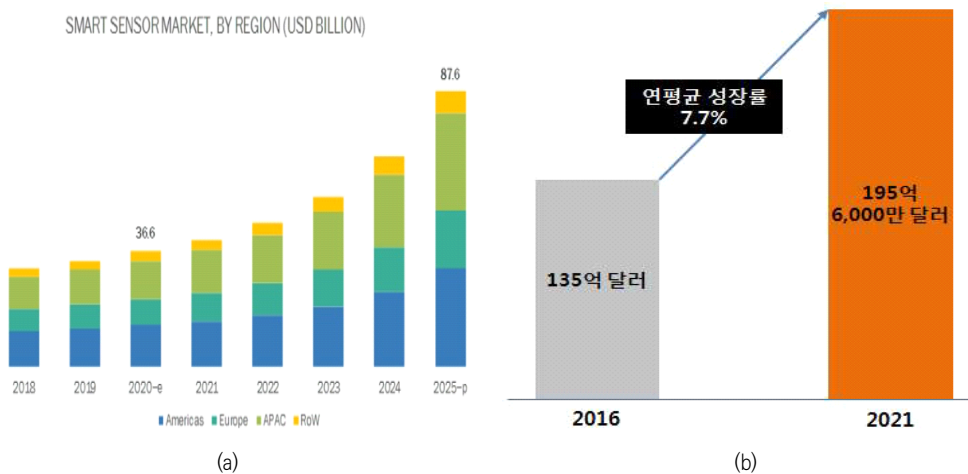
지능정보기술 및 위성영상 융합을 통하여 아프리카의 토양 유형, 토양수분 변화를 통한 가뭄 가능성 등을 나타내는 다양한 토양지도를 제작하는 프로젝트

<자료 >ETRI - 직독식 수질복합센서 및 초분광영상 기반 시공간 복합 인공지능 녹초 예측 기술 보고서 2019. 5, https://dbr.donga.com/article/view/1203/article_no/7252/ac/a_view

III. AI 기반 IoT 기술

1. AI 기반 IoT 기술 정책 동향

AI, IoT, 빅데이터가 모든 산업에서 성장 동력으로 주목받으면서 데이터 수집과 활용의 가치가 높아지고 있다. 이에 따라 압력, 온습도, 가속도 등 물리량을 전기적 신호나 데이터로 변환하는 센서의 수요가 증가하고 있는 추세이다.[8] 특히 세계 IoT 센서 시장에서 환경 관련 주요 센서의 비중은 2022년까지 53.7%로 전망하고 있다. 특히 온도센서, 조도센서, 환경 모니터링 센서는 2015년부터 2022년까지 연 평균 각각 25%, 27%, 그리고 26%씩 성장하였고 2022년 각각 147.2억 달러, 97.0억 달러, 그리고 57.4억 달러로 성장할 것으로 예측하고 있다.[9][그림3]에서 보듯 바이오센서 시장 점유율은 증가하는 추세를 보인다.



(자료)Marketsandmarkets, Environmental Monitoring Market, 2017, Smart Sensor Market, 2020.3

그림 3. 스마트센서 시장 (a)스마트 센서 시장 점유율, (b) 글로벌 환경 모니터링 센서 시장 규모

최근 국내 환경부는 ‘안전한 환경, 모두가 건강한 사회’를 비전으로 제 2차 환경보건종합계획(2021~2030년)을 발표하였다. 최근에는 환경오염(미세먼지 질량, 수농도, 독성물질 등)에 대한 국민의 관심이 증가하면서 이를 측정할 수 있는 스마트 센서에 대한 수요가 상당할 것으로 예상하고 있다.[10] 전 세계적으로 스마트센서 시장 내 환경 센서의 비중은 50% 이상이지만 현재까지 국내 환경 센서 관련 기업의

비중은 상대적으로 작고, 원격제어, 오작동 등 데이터의 신뢰성의 문제가 제기되고 있다. 따라서 환경부는 환경유해인자 사전예방관리에서부터 피해 대응 복구까지 범위가 확장하여 인공지능 기술을 접목한 환경보건 분야 기술의 고도화의 필요성을 언급하였다.

2. AI 기반 IoT 연구개발 기술 내용

‘ICT 기반 환경모니터링 센서 검증 플랫폼을 구축’ 연구과제는 환경 센서 무선 네트워크 신뢰성 시험규격(LoRa, NB-IoT, WiFi), 원격 모니터링 기반 신뢰성 검증(센서 빅데이터 구축, 신뢰성 진단 알고리즘 개발 등)에 따른 데모 시스템을 구축 중에 있다. 본 연구는 [그림 4]에서 볼 수 있듯 환경 센서 신뢰성 진단의 주어진 데이터를 기반으로 수행하였으며, 센서 신뢰성 진단의 특징을 비교한 후 알고리즘을 설계하였다.[11]

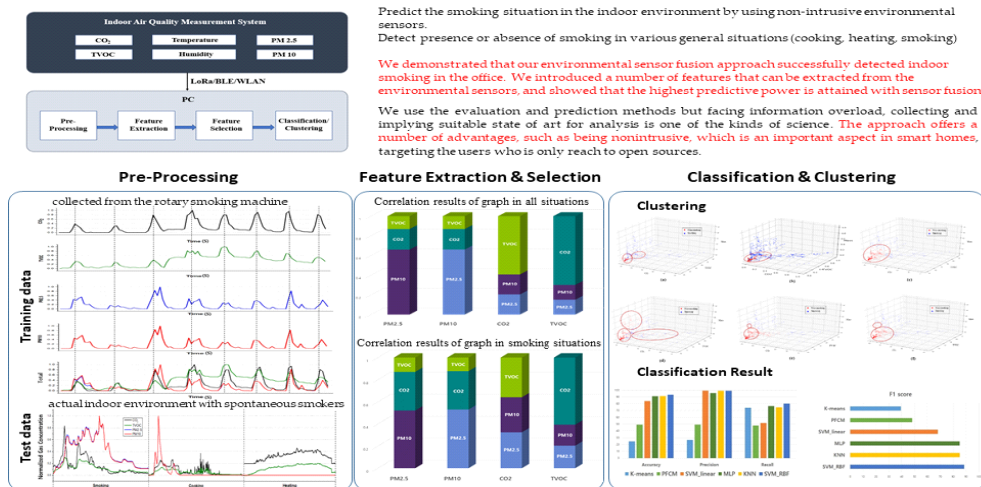


그림 4. 실내 공기질 악화(담배, 요리, 난방 상황 등) AI 전처리, 특징 도출 및 선택, 구분 및 군집화 방법

위 연구와 더불어 환경보건관련 된 지역기반 등 빅데이터 DB를 구축하여 AI를 적용한 감시체계를 고도화하는 연구를 진행 중에 있다.

빅데이터 플랫폼 기반 설계에 있어서 환경보건정보시스템과 문헌정보를 연계하여 데이터를 정형화 하고 구조화 한다. [그림 4]에서 보면 빅데이터를 구성하기 위해서

는 데이터 형태는 크게 정형데이터와 비정형 데이터로 구분한다. [그림 5]에서 정형 데이터는 다양한 IoT 기반 환경센서, 웨어러블센서 등 각종 디바이스와 환경보건 공공데이터(환경부, 기상청 정보 등)를 통해 수집한다. 비정형 데이터는 환경 측정자료에서 측정된 자료는 유형별로 분류하여 재구성하여 빅데이터를 활용한 수요, 예측 분석에 주로 사용할 수 있다. 특히 비정형 데이터의 정형화는 부분은 데이터소스 부터 수집해서 각 시스템별로 데이터를 구분을 하고 전사메타표준관리체계에 따라서 정리되는 사항(지침, 단어, 도메인 등)을 통해서 표준관리해서 정리할 예정이다. 또한 정리되지 않은 것들은 비표준 매핑을 통해서 유형별로 데이터 마이닝을 할 수 있다.

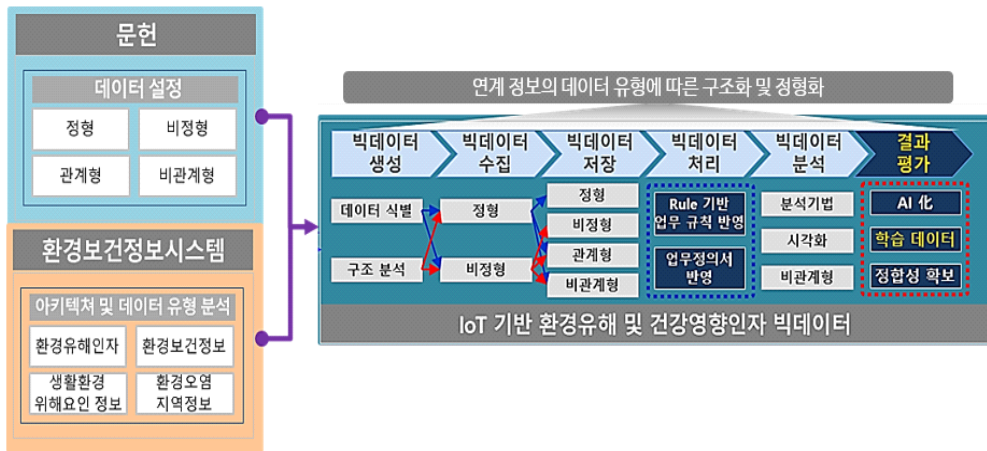


그림 5. 환경보건시스템 연계 및 빅데이터 플랫폼 설계 및 구축

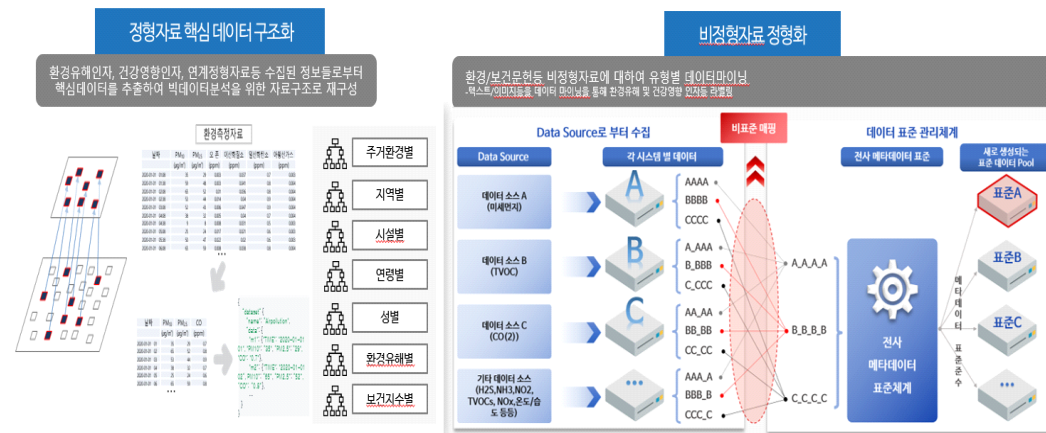


그림 6. 환경보건시스템 연계 및 빅데이터 플랫폼 설계 및 구축

[그림 7]과 같이 다양한 환경유해인자 및 건강영향인자의 원시 데이터를 원활하게 수집하기 위해서는 기존 디바이스 현황(하드웨어, 소프트웨어, 통신 방식, 저장체계 등) 및 통신방식별(블루투스, 와이파이, LTE, 5G 등) 특성 분석을 통해 초기데이터 수집용 디바이스 선정 및 IoT 모듈 설계하였다.

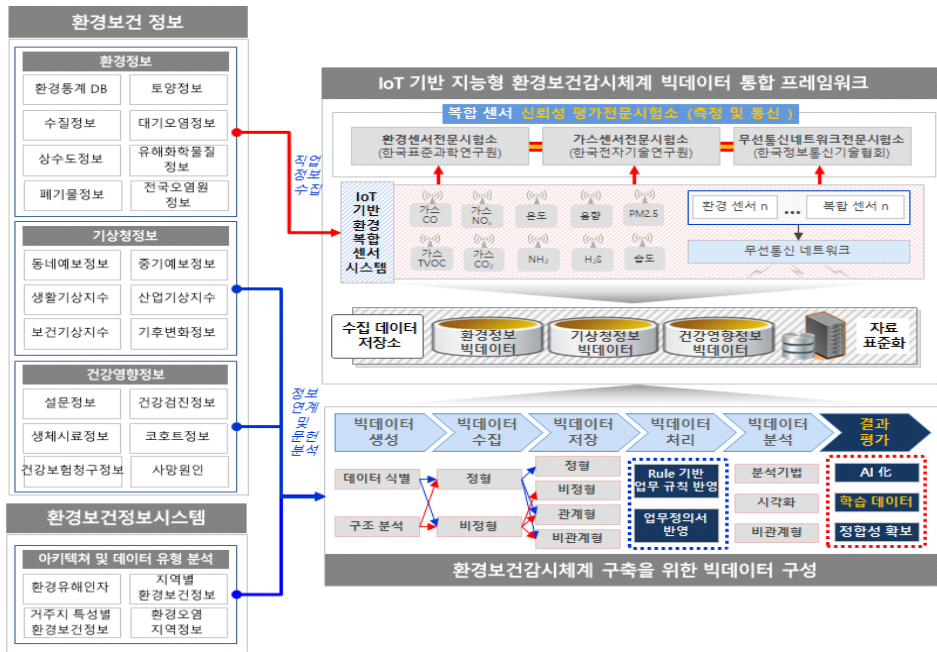


그림 7. 디바이스 선정 고려사항 및 IoT 모듈 설계

[그림 7]에서 볼 수 있듯 온·습도 센서의 경우 온도와 습도를 측정하는 감지 (Sensing)부와 감지부에서 온 신호를 디지털로 바꿔주는 아날로그부, 디지털 신호를 해석하는 연산부, 그리고 이를 무선 신호로 바꿔 중앙 클라우드나 관제 시스템으로 전송해주는 RF부로 구성되어 있으며 감지부에서 두 전극의 저항 차이를 알아내 아날로그 신호 처리 블록에 보내면 아날로그 부에서 이를 디지털로 변환해 연산부에 보내고, 연산부는 이를 처리해 RF부로 보내게 되며 RF부는 이 신호를 무선 신호로 바꿔 송출하게 된다. 측정된 센서 데이터의 신뢰성 확보하는 방법으로는 [그림 8]에서 보듯 첫 번째 빅데이터를 이용해서 원격으로 센서의 상태를 검사하고 센서의 이상상태를 검출 및 교정하기 위해서 필요한 데이터 구성요소를 정의하여 필요 데이터를 수집하고 센서의 설정을 변경 할 수 있는 방법에 대한 연구를 하였다. 이는 서버에서는

빅데이터 처리를 통하여 센서에서 수집된 데이터를 이용하여 센서 데이터의 오류를 검출하는 알고리즘 구현하였다.

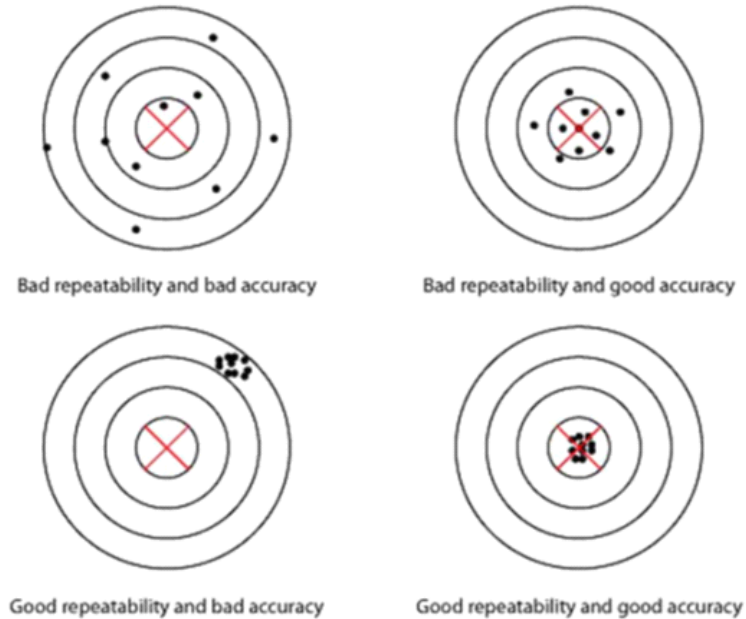


그림 8. 센서 데이터 오류 검출 방법 예시

두 번째로는 통계기반 기법으로 Z-값을 활용하였으며, 관측 값 추정을 위해 Kalman Fitter 등을 활용하였다. 또한 [그림 9]와 같이 통계기반으로 신뢰성 진단을 위해 탐색적 데이터 분석(EDA)를 활용하여 데이터의 구조적 특성을 알아낼 수 있었다.

환경유해인자 및 건강영향인자 빅데이터 활용 서비스 구현 기술 개발을 위해 인공지능 기반 환경유해인자 및 건강영향 데이터를 생산하고 이를 결측·이상 데이터 처리 알고리즘 등 검수 모듈 개발에 필요한 AI 구현 필요하다.

인공지능 분석 및 예측을 위해 실시간으로 수집된 데이터는 bigdata 저장소, file storage와 [그림 10]와 같은 기계학습(지도, 비지도, 강화학습)을 활용하여 [그림11]같은 방법으로 F1 score, ROC를 통한 모델을 평가하여 AI를 이용한 데이터를 분석할 수 있다.

센서 신뢰성 진단을 위한 알고리즘 설계

일반상황에서의 센서데이터 세부 분석

- 담배에서 발생하는 유해가스 중 TVOC의 세부 분석과 예측을 통해 실제상황에서의 알람과 같은 상황에 이용하고자함.

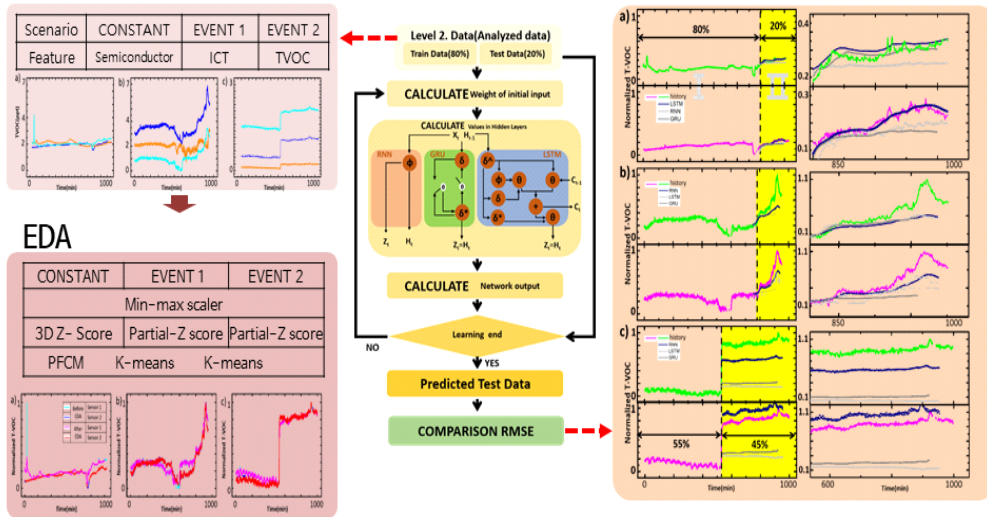


그림 9. 센서 데이터 세부 분석 절차

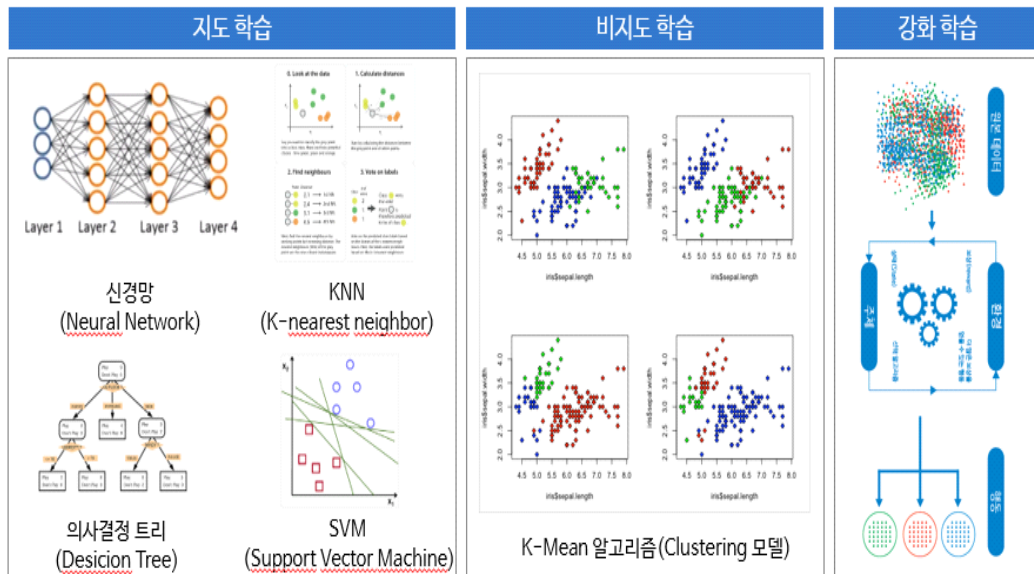


그림 10. 센서 데이터를 적용할 기계학습(지도, 비지도, 강화학습)

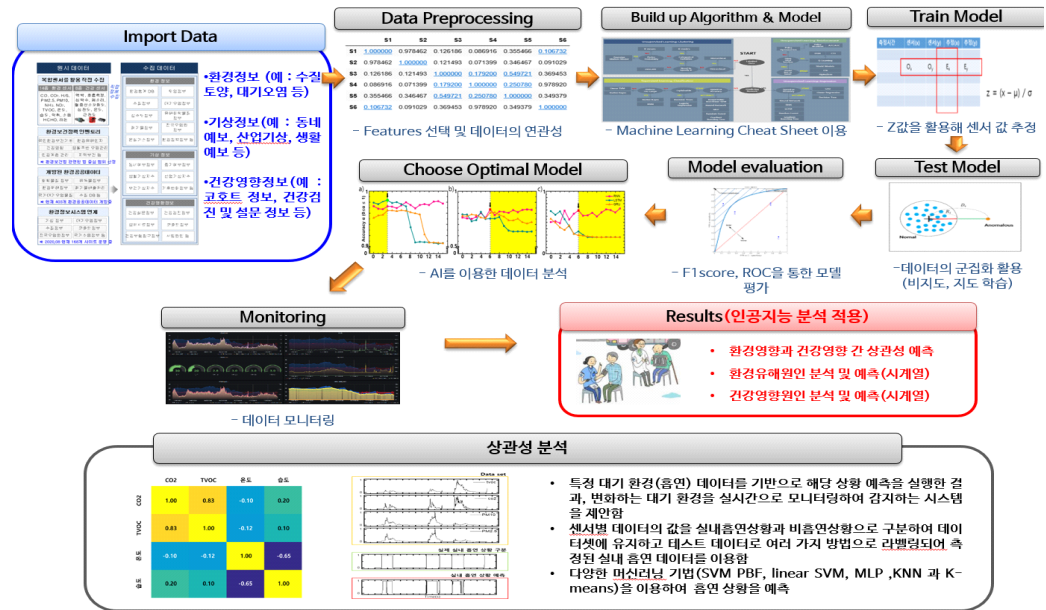


그림 11. 인공지능 분석 및 예측

이를 내부적, 외부적으로 쉽게 사용할 수 있도록 OPEN API로 개발하여 GIS 기반의 시각화할 수 있다.

IV. 결론 및 시사점

최근 수년간 진화를 거듭하고 있는 인공지능(AI) 기술은 4차 산업혁명을 선도하는 주요 차세대 기술들 가운데 가장 혁신적인 기술로, 산업뿐 아니라 인간의 삶 자체를 변화시키는 핵심 역할을 수행 할 것으로 보여진다. 최근 5G시대를 맞아 IoT 디바이스 보급으로 인해 데이터가 폭증함에 따라 빅데이터는 더욱 중요해 질것으로 예상되고 있는 실정이다. 특히 환경오염문제가 대두되면서 국가차원에서 환경 난제 예측 분석을 위해 기상, 환경, 교통, 생산 등 다분야 빅 데이터 분석 R&D 추진 중에 있으며 많은 데이터를 확보하기 위해서 지속적인 환경측정센서 개발이 필요하며 센서의 신뢰성을 고도화하기 위해 본 고에서는 센서 데이터를 통해 환경유해인자와 건강영향인자간의 상관관계를 파악하고, AI를 이용하여 핵심데이터를 구조화 및 정합성을 설계하여 기존 선행연구와 차별성을 두었다. 이는 환경유해인자를 측정할 센서에 대한

물리량 측정 신뢰성, 센서 통신 신뢰성, 센서 원격검증 기반센서 데이터 신뢰성에 대한 평가·검증체계를 구축하여 센서 데이터의 신뢰성을 확보하여 다각적인 활용이 가능하다. 그러나 아직도 우리나라는 세계 글로벌 기업들과 비교해서 상대적으로 인공지능 기술 역량과 데이터 측면에서 상당한 격차를 가지고 있다. 이를 위해 국가적 차원에서 부처 및 연구소간의 협업을 강화하여 인공지능기반 다분야의 R&D 및 인프라 자원을 위한 중장기적 정책방안이 필요하다. 두 번째로는 사물인터넷, 스마트시티, 스마트 자동차 등 미래 성장분야와 인공지능 적용 산업에 세 번째는 IoT 기반의 빅데이터 확보하기 위해서 개인정보보호 환경이 구축되어야한다. 많은 데이터로 인해 개인의 프라이버시 침해 등 인권침해로 인해 우려할만한 부정적 요소가 존재한다. 국내 법체계와 해외 입법 동향을 분석한 결과를 반영하여 글로벌 동향과 조화를 이루면서 우리의 실정에 맞는 법제 정비(안)을 마련해야한다.

따라서 이는 취약지역, 집단별 불평등을 해소할 수 있는 환경보건 안전망 강화와 다양한 수요를 충족할 수 있어 맞춤형 환경보건정책 및 서비스에 활용이 가능하다. 또한 환경보건 관련 기술 개발로 신규 환경유해인자에 대한 선제적 예측 및 관리 등이 가능하여 국민 삶의 질 향상에 기여할 수 있다.

[참고문헌]

- [1] 조재혁, 나영식, “인공지능(SW)”, 한국과학기술기획평가원(KISTEP), 2018-16호
- [2] “2021 당신이 알아야 할 5가지 주요기술”, Insight campus, 2020. 9
- [3] <https://www.gartner.com/en>, “2021 AI 9대 기술전략”, 2020.12
- [4] IRS Global, “2021 글로벌 AI, 빅데이터 관련 시장 사업화 동향과 기술개발 전략”, 2020-18호
- [5] 데이비드 스카츠키 외 2인 “지능형 IoT : AI의 힘을 사물인터넷과 결합하기”, Deloitte Korea Review, No.10
- [6] 정보통신정책연구원, ‘인공지능, 코로나 19를 만나다.: 코로나19로 되짚어보는 AI의 현재와 미래’, 2020, 4
- [7] 과학기술정보통신부, “인공지능(AI) R&D 전략”, 2018.5
- [8] https://www.e4ds.com/sub_view.asp?ch=2&t=0&idx=12188, “
- [9] 간순영, 윤성지, 신용승, “빅데이터를 활용한 환경보건서비스에 관한 기초연구”, 한국환경정책평가연구원, 2015-07.
- [10] <http://www.aitimes.com/news/articleView.html?idxno=123322>, “대기오염에 적용되는 AI 기술”, 2019.12
- [11] “ICT 기반 환경모니터링 센서 검증 플랫폼을 구축”, IITP, 송실대학교