

## 빌딩 운영(Building Operation) 및 유지관리(FM)에서의 머신러닝(Machine Learning) 적용에 관한 최근 연구동향 및 활성화 방안



홍성일

Ph.D. Candidate, Georgia Institute of Technology, si01.hong@gatech.edu

저자는 조지아 공대(Georgia Institute of Technology)에서 건축공학 박사과정 중이며 부전공으로 컴퓨터 공학을 하였다. 현재 최종 논문 심사만을 남겨두고 있다. 박사과정을 시작하기 전 스타트업에서 센서와 인터넷 정보를 활용한 스마트 홈 프로젝트를 진행하였으며 삼성물산 건설부문 소속으로 삼성전자 수원 주차타워 현장에서 건축기사와 충남 삼성고등학교 현장에서 공무 주임으로 근무하였다. 박사과정 연구 주제로 머신러닝과 텍스트 마이닝 기술을 활용한 스마트 빌딩(smart building) 구현과 사용자 중심의 업무 공간 운영(workplace management) 연구를 진행 중이다.

### 1. 서론

전체 에너지관련 이산화탄소 배출 중 28%가 빌딩 운영 부문이 차지하고 있는 가운데, 세계적으로 탄소 배출 저감이 강조되고 있어 에너지 사용량 감소를 위한 빌딩 디자인과 효율적인 빌딩 운영의 중요성이 대두되고 있다(Abergel et al., 2020). 하지만, 에너지 사용 저감 달성만이 아닌 빌딩 사용자의 만족도 충족시켜야 하기에, 데이터를 기반으로 보다 에너지 효율적이면서 사용자 경험에도 효과적인 운영을 위한 의사결정이 요구된다. 빌딩 운영 및 관리 과정에서 다양한 종류의 데이터가 발생함에도 불구하고, 의학, 제조업, 리테일과 같은 다른 학문 혹은 산업분야에 비해 유지관리(FM) 산업에서는 머신러닝(machine learning) 기술 활용이 활발하게 이루어지지 않는 실정이다. 머신러닝 기술은 기존의 데이터를 기반으로 수치 예측부터 프로세스 자동화까지 광범위하게 활용이 가능하므로 빌딩 관리 및 운영시 해당 기술의 적용이 필수적이다. 본 고에서는 빌딩 운영을 하는 과정에서

발생하는 데이터 기반의 의사결정을 위해 머신러닝이 어떻게 적용될 수 있는지에 대한 최근 연구동향과 함께 활성화 방안까지 살펴보고자 한다.

### 2. 본문

#### 2.1 연구동향

##### 1) 머신러닝 개요

사전에 주어진 데이터를 기반으로 하는 머신러닝은 크게 두 가지로 분류된다(Bishop, 2006). 한 가지 방법은 예측하고자 하는 변수와 관련된 데이터의 입력값과 출력값 혹은 분류값(label)이 제공되어, 주어진 데이터를 학습한 머신러닝 모델을 생성하여 새로운 조건이 입력되었을 때 특정 값이나 클래스를 예측하는 지도학습(supervised learning)이다. 다른 한가지로는 정답(label)을 알려주지 않아 지도학습이 불가능한 데이터를 비슷한 성격의 데이터끼리 묶어 군집화(clustering)하는 비지도학습(unsupervised learning)이 있다. 예를 들어, 한 지역의 주택 가격정보들을 기반으로 그 지역의 새로 올라온 주택의 가격을 예측하거나 강아지와 고양이 사진을 학습한 후 새로운 사진이 강아지인지 고양이인지 예측할 때 사용하는 방식이 지도학습이며, 카테고리 정보(label)가 없이 비슷한 생김새와 색깔의 과일 이미지를 분류하는 것과 같이 비슷한 성격의 데이터 값들을 분류할 때 사용하는 것이 비지도학습이다. 또한 머신러닝 모델을 생성할 때는 알고리즘 별로 사용되는 데이터의 특성에 따라 다른 수준의 수행능력을 보여주므로 여러 알고리즘을 테스트하고 최적합인 알고리즘을 찾아서 사용해야 한다.

## 2) 센서 기반 데이터와 머신러닝의 적용

유지관리(FM) 영역의 데이터로는 건물 설비를 제어하고 에너지 사용을 관리하는 시스템인 Building Management System (BMS)/ Building Automation System (BAS)/ Building Energy Management System (BEMS)에서 수집되는 센서 데이터와 빌딩의 유지보수 및 관리 활동 프로그램인 Computerized Maintenance Management System (CMMS), Computer Aided Facility Management (CAFM), Integrated Workplace Management System (IWMS)에서 수집되는 사용자 입력 데이터 등이 있다. 위의 수집된 데이터는 전문 지식을 통해 예측 및 최적화하고자 하는 변수와 그와 관련된 필요 데이터들이 선택된다. 그 후, 코스트(손실) 함수를 결정하고 적합한 머신러닝 알고리즘을 선택하여 학습시킨다. 머신러닝 모델의 성능을 분석하고 개선하기 위해서는 머신러닝 관련 지식 뿐만 아니라 전문 지식이 요구된다.

예를 들어, BMS/BAS/BEMS는 공조 설비(Heating, Ventilation, and Air Conditioning (HVAC)), 전기, 조명 등을 제어하는 프로그램으로 빌딩의 효율적인 제어를 위해서 인체 감지 센서, 온·습도 센서, 이산화탄소 센서, 에너지(전기) 소비량 계량기 정보 등을 활용한다. 빌딩의 에너지 사용량과 빌딩의 크기, 층수, 사용 목적 등의 빌딩 정보를 수집하여 새로운 빌딩을 디자인 할 때 해당 빌딩의 에너지 사용량을 예측해 볼 수 있다(Duncan et al., 2014). 다른 예로, 빌딩을 운영하면서 대부분의 빌딩들이 시간별로 냉·난방 가동 시간이 사전 설정되어 있어 과도하게 냉·난방 설비가 작동하거나 필요한 시기에 작동하지 않아 사용자의 컴플레인을 받는 경우가 빈번하다. 이런 경우 사용자의 만족도도 충족시키지 못하고 불필요한 에너지 사용까지 발생하는데 이를 개선하기 위해 인체 감지 센서 정보, 실내 온도, 이산화탄소 센서 정보를 수집하여 실(room) 별 사용 패턴을 분석하고 앞으로의 사용을 예측하여 냉·난방 가동 시간의 재조정이 머신러닝을 통해 가능하다(Mamidi et al., 2012). 이는 에너지 사용량 감소와 함께 사용자 만족도를 높이는 효과가 있다. 현재까지 에너지 사용과 관련하여 전통적인 방식의 시뮬레이션과 회귀분석 등의 분석방법이 있었지만 머신러닝 알고리즘의 장점은 다양한 데이터와 기법을 통하여 에너지 사용량 예측의 정확도를 획기적으로 높이는데 있다 (Rai et al., 2019). 이는 빌딩 에너지 사용 최적화 측면에서 머신러닝 적용의 필요성을 시사한다.

## 3) 유저입력 기반 데이터와 머신러닝의 적용

센서에 의해 수집되는 데이터 외에도 유지관리 프로그램을 사용하는 유저의 활동과 비용 등이 기록되는 CMMS와 IWMS 데이터 또한 의미 있게 사용될 수 있다. CMMS와 IWMS 소프트웨어를 통해 공간, 시설 운용 및 보수 등의 유지관리 활동 그리고 예산까지 다양하게 관리가 가능하다. 이러한 소프트웨어에서 수집되는 유지보수 데이터, 에너지 사용 데이터, 그리고 빌딩 정보를 활용하여 빌딩의 생애 주기 비용(life cycle cost) 분석이 진행되었다(Gao et al., 2019). 머신러닝의 알고리즘 중 하나인 다층 퍼셉트론(multi-layer perceptron)을 기반으로 하는 회귀모델이 사용되었으며, 건물의 연면적(gross square feet), 냉·난방 에너지 사용량, 빌딩 사용목적, 그리고 사용 연한이 생애 주기 비용에 기여하는 주요 인자로 밝혀졌다. 하지만, 이렇게 숫자로 이루어진 정보들은 머신러닝 알고리즘을 적용하기가 용이하여 현재까지 많은 연구가 이루어져왔으나 유지보수 활동이 기록된 정보들은 문자 정보로 비구조화 되어 있고 사용자가 입력하는 과정에서 카테고리 정보 미기입 등의 문제로 데이터가 불완전하여 방대한 데이터가 기록되어 있음에도 불구하고 현재까지 잘 활용되어 오지 않았다(Ahmed et al., 2017).

필자는 유지보수 프로세스를 단순화하고 기존에 존재하는 데이터의 활용 방안을 제시하고자 CMMS를 통해 기록된 유지보수 활동 데이터를 텍스트 마이닝과 머신러닝 기술을 활용하여 각 활동에 해당하는 이슈 카테고리 코드를 자동으로 입력하는 모델을 개발하였다(그림 1) (Hong et al., 2020). 데이터의 분류는 분석과정을 빠르고 정확하게 할 수 있도록 도와주기 때문에 필수적이다. 예를 들어, 고객센터 센터에 접수되는 컴플레인을 자동으로 분류하여 긴급한 사안은 빠르게 처리하고 영역별로 분류된 컴플레인을 분석함으로써 서비스의 질(quality)을 개선할 수 있기 때문에, 서비스 제공자 입장에서 수집된 문자 데이터의 분류는 중요하며 이를 위해 다양한 연구가 진행되어 왔다(Hunt et al., 2005; Pyon et al., 2011). 빌딩 사용자의 컴플레인 또한 문자 데이터로 접수·기록되기 때문에 룰 기반(rule based)의 자동화 방식이 작동되지 않고, 동시에 빌딩 내의 다양한 시설과 수많은 보수 관련 카테고리의 존재로 업무를 처리하는 근로자(프로그램 사용자)에 의해서도 입력이 원활하게 이루어지지 않고 있다. 이 같은 이유로 텍스트를 기반으로 구성된 데이터의 카테고리를 자동으로 분류하는 시스템의 필요성을 전문가

WO #	WO Desc	Type	Problem Code	Craft	Facility
51202-2019	ISYE MAIN ROOM 202 TOO WARM THERE IS A LOCKING DOOR CONCEALING A CUT OFF VALVE THAT HAS BEEN TORN FROM THE WALL AND SINK	P	05A		19 ISYE (ANNEX)
51529-2019	IS STOPPED UP	P			1 MRDC
51529-2019	THERE IS A LOCKING DOOR CONCEALING A CUT OFF VALVE THAT HAS BEEN TORN FROM THE WALL AND SINK IS STOPPED UP	P			4 MRDC
51536-2019	REPAIR DISH RM FAUCET 53940135102 WEST VILLAGE PLEASE ISSUE WORK ORDER TO REPAIR SINK IN DISH ROOM AT WEST VILLAGE. THE SINK WILL NOT SHUT OFF.	P	04B		4 W CAMPUS DINING
51538-2019	REPLACE CEILING TILE IN TECH REC	P	10B		10 STU CTR - REC
51563-2019	ISYE MAIN-REPLACE CEILING TILES IN ROOM 448 AND 425D	P	10B		10 ISYE (ANNEX)
51600-2019	LOVE BUILDING CLOGGED URINAL IN 2ND FLOOR MSE SIDE PLEASE UNCLOG. THERE'S CURRENTLY A GRAY TRASH BAG OVER THE URINAL.	P	04A		4 MRDC II

➔

	Description	Label
Data #1	ISYE MAIN ROOM TOO WARM	05A (Too hot)
Data #2	IBB ROOM LEAK IN CEILING REFER TO EMAIL PHOTO	05C (AC Leak)
Data #3	EBB ROOM TOO COLD	05B (Too Cold)
Data #4 (New)	IBB PRL WATER LEAK EMERGENCY	? -> 05C

그림 1. 머신러닝 기반 빌딩 사용자 컴플레인 분류 자동화 예시 (그림: 홍성일)

들과의 인터뷰를 통해 확인할 수 있었다. 이 연구를 통해 소프트웨어 사용자의 입력 과정 중 쉽게 발생하는 실수의 패턴과 카테고리 재분류의 필요성이 대두되었으며, 향후 데이터 분석의 정확도 향상과 유지보수 프로세스 및 서비스 질 개선을 위해 카테고리 정보 입력의 자동화 그리고 기존의 미사용 데이터 활용의 필요성 또한 강조되었다.

## 2.2 활성화 방안

연구 동향 분석을 통해 빌딩 운영시 발생하는 다양한 데이터를 머신러닝 알고리즘을 적용한 에너지 사용량 예측 및 절감, 유지보수 프로세스의 단순화, 생애주기 비용 예측까지 여러 활용방안을 살펴보았다. 하지만, 그 성능과 효과에도 불구하고 유지관리(FM) 영역에서 머신러닝이 잘 활용되지 않고 있다. 따라서, 머신러닝 적용의 활성화를 위해 요구되는 노력에 대해서 논의해 보고자 한다.

첫째, 가용한 데이터에 대한 이해와 머신러닝 적용에 필요한 데이터의 정의가 필요하다. 앞서 언급했듯이 빌딩 관리 및 운영 프로그램과 데이터의 종류가 다양하기 때문에 어떤 데이터가 생성되고 목적을 달성하기 위하여 필요한 데이터가 무엇인지 구체화하기 위해 실무 전문가와 머신러닝 엔지니어의 협의가 필요하다. 둘째, 머신러닝 기술을 적용하기 위해서는 데이터를 해석하고 머신러닝으로 해결할 수 있는 문제를 정의하는 능력이 필수적이다. 이를 위해서는 전문 영역의 기반 지식 뿐만 아니라 머신러닝의 기초 지식을 갖추고 있어야 가능한데 현재 해당 지식을 가지고 있는 인력이 부족하다. 따라서, 기반 지식을 갖춘 실무 전문가들이 머신러닝 혹은 인공지능 관련 기초 교육을 받을 수 있는 환경조성과 함께 머신러닝 기초지식을 통한 인공지능 엔지니어들과의 접점을 만드는 것이 중요하다. 셋째, 머신러닝 알고리즘의 사용성 이슈로 데이터를 전처리(pre-processing) 하는

과정과 알고리즘 적용 또한 코딩 언어가 아닌 앱과 같은 단순화된 툴이 필요하다. 이를 통해 일반 사용자들의 머신러닝 기술에 대한 접근성이 개선되어야 더 활발하게 적용될 수 있을 것으로 예측된다. 이는 IBM, 구글, 마이크로소프트 등의 주요 테크기업들에 의해서 빠르게 개발 중이며 상용화 추이와 함께 사용 방안이 계획되어야 한다.

## 3. 결론

머신러닝 기술은 컴퓨터 공학영역에서의 발전 속도에 비해서 산업에서의 적용 속도가 많이 뒤쳐져 있으나, 점차 활성화되고 있다. 특히, 건설 및 유지관리 영역은 보수적인 성격의 산업 특성상 그 적용이 다른 산업군에 비해 더딘 것이 사실이다. 이를 극복하기 위해서 결정권자 및 관리자들의 관심이 요구된다. 특히, 전문가들이 인공지능 기술을 활용한 툴을 직접 개발하는 것이 아닌 기초지식 교육을 통해 전반적인 이해를 기반으로 인공지능 엔지니어들과의 원활한 의사소통이 보다 효율적이고 효과적인 스마트 빌딩(smart building) 디자인과 운영의 발전에 기여할 수 있다고 믿는다. 이번 기고를 통하여 머신러닝과 그 적용에 대한 소개가 잘 이루어졌길 바라며, 한국의 건설 및 유지관리 산업의 머신러닝 기술적용과 발전에 조금이나마 보탬이 되었으면 한다.

## 참고문헌

- Abergel, T., Delmastro, C., & Lane, K. (2020). *Tracking Buildings 2020*. International Energy Agency (IEA). <https://www.iea.org/reports/tracking-buildings-2020>
- Ahmed, V., Tezel, A., Aziz, Z., & Sibley, M. (2017). The future of big data in facilities management: opportunities and challenges. *Facilities*. <https://doi.org/https://doi.org/10.1108/>

f-06-2016-0064

3. Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
4. Duncan, S., Balchanos, M., Sung, W., Kim, J., Li, Y., Issac, Y., & Mavris, D. (2014). Towards a Data Calibrated, Simulation-Based Campus Energy Analysis Environment for Situational Awareness and Future Energy System Planning. *Proceedings of the ASME 2014 8th International Conference on Energy Sustainability*, 1-10. <https://doi.org/10.1115/ES2014-6695>
5. Gao, X., Pishdad-Bozorgi, P., Shelden, D. R., & Shu, T. (2019). Machine Learning-based Life-cycle Cost Analysis for Educational Facilities. *55th ASC Annual International Conference*, October, 393-400. [https://www.researchgate.net/publication/334262496\\_Machine\\_Learning-based\\_Life-cycle\\_Cost\\_Analysis\\_for\\_Educational\\_Facilities](https://www.researchgate.net/publication/334262496_Machine_Learning-based_Life-cycle_Cost_Analysis_for_Educational_Facilities)
6. Hong, S., Kim, J., & Yang, E. (2020). *Automated Text Classification of Historical Maintenance Data of Higher Education Buildings using Text Mining and Machine Learning Methods*. Manuscript Submitted for Publication.
7. Hunt, R. A., Killen, M. C. P., Bosch, V. G., & Enríquez, F. T. (2005). TQM and QFD: exploiting a customer complaint management system. *International Journal of Quality & Reliability Management*. <https://doi.org/10.1108/02656710510572977>
8. Mamidi, S., Chang, Y.-H., & Maheswaran, R. (2012). Improving building energy efficiency with a network of sensing, learning and prediction agents. *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, 45-52.
9. Pyon, C. U., Woo, J. Y., & Park, S. C. (2011). Service improvement by business process management using customer complaints in financial service industry. *Expert Systems with Applications*, 38(4), 3267-3279. <https://doi.org/10.1016/j.eswa.2010.08.112>
10. Rai, P., Nassif, N., Eaton, K., & Rodrigues, A. (2019). Applications of Machine Learning in Building Energy Prediction and Savings. *Energy*, 1, 1-2. <https://doi.org/10.3844/erjsp.2019>