

BEMS 시스템 구축을 위한 클러스터링 및 머신러닝 기반 건물별 에너지 소모량 예측 모델

[IE빅데이터분석][에너지와사회] 강의 페어링, 이진희 교수님 지도
산업공학과_201920209_박찬영

연구 목적

현재 IT 기술을 접목한 건물의 에너지 효율을 관리하는 건물에너지관리시스템(BEMS)에 대한 관심이 늘고 있는 상황이다. BEMS에 대해 더 나은 이해를 하고자 정확한 건물 에너지 소비 예측 모델을 제작한다. 클러스터링을 통해 건물 전력 소요량에 유의미한 요인을 찾고 머신러닝 알고리즘을 통해 에너지 소비 패턴이 다른 각각의 건물들에 대한 에너지 소비 예측 모델을 만들어 건물별 에너지 소비량을 예측할 것이다.

[IE빅데이터분석]에서의 강의 페어링

- 건물 전력 소요량에 유의미한 요인이 무엇인지 분석.
- 유의미한 변수들을 통해 정확한 모델을 제작.
- 이러한 부분들을 [IE빅데이터분석]에서 배운 방법론들을 적용.

데이터 소개

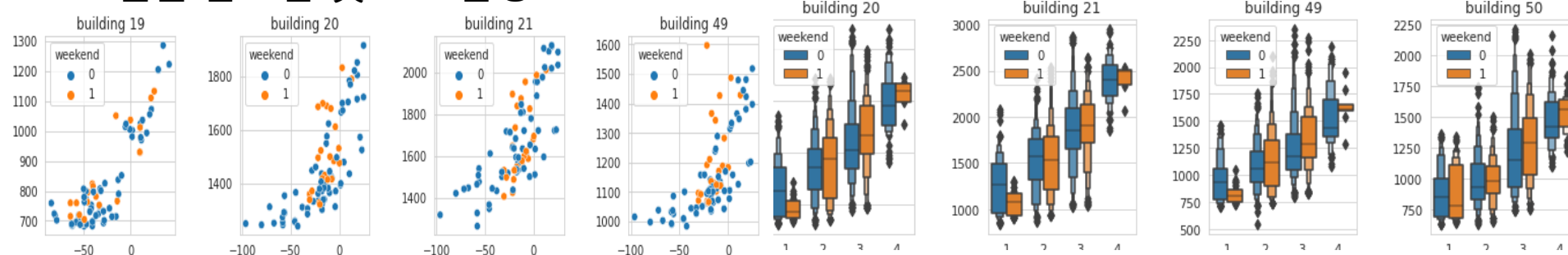
- 6월 1일부터 8월 24일 까지, 1시간 단위 전력사용량을 포함하는 기상 관측 데이터.
- 컬럼 : 건물번호, 시간, 전력사용량, 기온, 풍속, 습도, 강수량, 일조, 비전기냉방설비운영, 태양광보유.
- 사람들의 행위의 중요한 기준이 되는 시간 요인, 사람의 행위에 영향을 주는 기상 요인들을 의미.

변수 선정



<회귀모델 선정>

- 기존에 에너지 데이터 : 자연적 요인에 영향을 받기 때문에 계절, 시간들에 따른 추세가 존재.
- 추세를 반영할 수 있는 시계열 도메인 예측 모델을 구상.
- 각 빌딩들의 전력에너지 사용량 패턴, 평균과 왜도의 측면에서 건물마다 매우 다른 것을 알 수 있음.
- 심지어 한 건물에서도 같은 요일임에도 불구하고 전력량의 차이가 큰 것이 관측.
- 회귀모델로 제작 및 각 건물별로 모델을 만드는 것으로 결정.



<유의미한 요인>

- 왼쪽은 냉방도일(냉방에 필요한 연료 소비량에 관한 기후지수), 오른쪽은 불쾌지수와 전력소모량 간의 각각의 그래프.
- 두 변수가 온도보다 더 설명력이 좋은 것으로 확인.
- 냉방도일과 불쾌지수가 전력소모량의 좋은 설명력.
- 일시적인 온도보다는 평균이므로 좀더 장기적인 성격이기 때문.
- 이러한 현상을 실제 상황에 대입하면 다음과 같다.
- : 빌딩에 있는 사람이 온도가 두시간 동안 31도에서 29도로 감소해도 에어컨의 온도를 높이거나 끄지 않는다.

[에너지와사회]에서의 강의 페어링

- [에너지와사회]에서 배운 전력과 밀접한 요인들을 참고.
- 변수의 유의미성을 분석할시 인과관계를 해석.
- 에너지 데이터에 대한 전반적인 이해 도움.

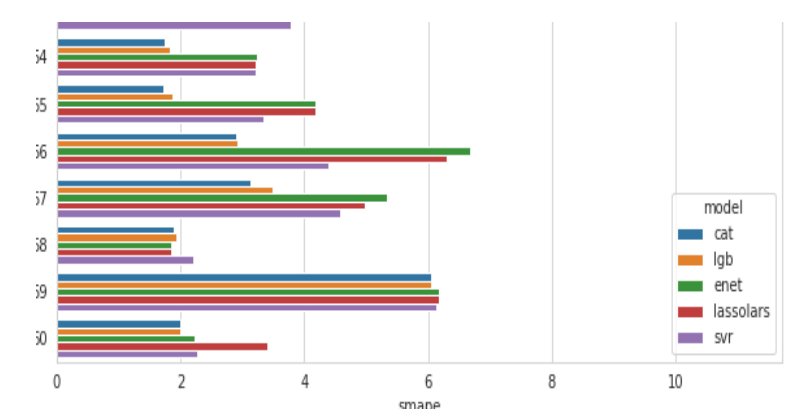
평가지표

<SMAPE>

에너지 소비 예측 시스템 특성상 실제값보다 적은 에너지 소비량을 예측할 시 정전, 전력부족 등 위험도가 큼. 그러므로 과대추정보다 과소추정에 더 큰 패널티가 부여 되는 SMAPE 오차가 더욱 효율적인 평가 지표임.

$$SMAPE = \frac{100}{n} \times \sum_{i=1}^n \frac{|Y_i - \hat{Y}_i|}{(|Y_i| + |\hat{Y}_i|) / 2}$$

예측 모델 제작



- CAT, LGB, ENET, LASSOLARS, SVR 모델들을 대상으로 GRID SEARCH 방식으로 최적의 모델 및 최적의 파라미터를 찾아 SMAPE 값을 비교.
- 예상대로 각 건물들마다 최적의 모델이 다름.
- 하지만 같은 클러스터끼리 있는 건물들끼리 같은 모델인 것을 확인

결론 및 고찰

- 건물의 전력 소모량을 예측하기 위해서 어떤 요인, 어떤 모델 선정에 대한 많은 고민을 했다.
- 그리고 많은 가설들이 틀리면서 다음과 같은 결론 도출.
- 1. 예상과 달리 온도의 설명력이 높지 않음.
- 2. 건물들의 소모량을 예측
- 전력 소모량에 추세가 존재하지 않아 회귀 모델을 고려.
- 3. 일시적인 성격의 요인보다는 비교적 장기적인 변수의 설명력이 더욱 좋음.
- 이렇듯 프로젝트를 진행하면서 당연시하게 생각했던 부분들이 틀렸다는 것을 알게 되었다.
- [IE빅데이터분석]에서는 단순히 방법론과 알고리즘을 배웠으나 이번 프로젝트를 통해 데이터 분석과 알고리즘을 진행할 때 데이터 도메인에 대해 충분한 이해가 기반이 되어야 한다는 것이 중요하다는 것을 깨달았다.