

SVR 알고리즘 기반 식품공장 에너지 소비 예측 모델의 성능 비교

이형아, 김동주, 구재희

고등기술연구원 에너지환경IT융합그룹

lhaeve@iae.re.kr, kimdongju@iae.re.kr, jaehoi@iae.re.kr

Comparison of Performance of Food Factory Energy Consumption Prediction Model based on SVR Algorithm

Hyungah Lee, Dongju Kim, Jae-Hoi Gu

Institute for Advanced Engineering Energy Environment IT Convergence Group

요약

본 논문은 FEMS에 적용하기 위한 머신러닝 기반 에너지 소비 예측 모델 연구의 일환이다. SVR 알고리즘을 기반으로, 대표적으로 사용되는 커널함수 중 Linear, RBF, Polynomial을 적용하였을 때의 예측 성능을 비교하였다. 연구 결과, 세 가지 모델 중 RBF 커널함수를 사용한 SVR 모델의 예측 정확도가 가장 높았다. 데이터의 추가적 확보와 파라미터 조정을 통해 예측 정확도를 높일 수 있을 것이며, SVR을 포함한 다양한 알고리즘을 활용하여 대상 식품공장의 에너지 소비 예측을 위한 최적의 모델 도출이 가능할 것으로 판단된다.

I. 서론

다양한 부문에서 에너지 소비 절감을 위해 기존의 수동적 에너지 관리 방식에서 나아가 능동적 에너지 관리 방식을 적용하고 있다. 이러한 측면에서 다양한 ICT(Information and Communication Technology) 기술을 융합해 실시간으로 에너지 소비를 모니터링하고 제어할 수 있는 에너지관리시스템(Energy Management System, EMS)에 대한 관심이 증대되고 있다[1]. 그 중 공장에 설치되는 EMS를 칭하는 ‘공장에너지관리시스템(Factory Energy Management System)’은 제3차 에너지기본계획을 통해 2025년부터 10만TOE 이상의 에너지 다소비 사업장을 대상으로 적용 의무화될 것으로 발표되었으며, 보다 효율적인 공장 에너지 관리를 위해 많은 사업장에서 FEMS의 도입을 고려하고 있다[2]. FEMS에는 공장 에너지의 실시간 모니터링, 소비 패턴 분석 및 예측, 에너지 소비 설비 및 유틸리티 제어 등의 다양한 기능이 포함된다. 그 중 에너지 소비 예측 기능은 에너지 소비 현황 및 패턴 분석과 더불어, 공장의 에너지 공급량과 비용을 적절하게 설계하기 위한 첫 단계라고 볼 수 있다.

본 논문은 FEMS에 적용하기 위한 머신러닝 기반 에너지 소비 예측 모델 선정 연구의 일환으로, 다양한 알고리즘을 활용해 식품공장의 에너지 소비 예측 모델을 구현하고 그 중 가장 높은 예측 정확도를 갖는 모델을 선정함에 목적을 둔다. 이를 위해 커널 함수별 서포트 벡터 회귀(Support Vector Regression)기반의 전력 에너지 소비 예측 모델의 예측 정확도를 비교하고자 한다.

II. 본론

본 논문은 대상 식품공장의 데이터 확보, 상관관계 분석을 통한 에너지 소비 예측 모델에 활용할 데이터 선정, 에너지 소비 예측 모델 구현, 에너지 소비 예측 모델 평가의 4단계를 포함한다.

II-1 대상 식품공장의 데이터 확보

본 논문의 연구범위인 전력 사용량 예측을 위해 2018년 1월 1일부터 2020년 12월 31일까지의 일(Daily) 단위 데이터를 활용했다. 활용된 데이터에는 (i)당일 및 전 일 전력 사용량, (ii)전체 제품 생산량, 1공장 및 2공장 제품 생산량, (iii)전체 인력 투입시간, 1공장 및 2공장 인력 투입시간, (iv)전체 투입인원 수, 1공장 및 2공장 투입인원 수, (v)외기온도, (vi)외기습도가 포함된다. 전력 사용량 데이터는 대상 식품공장의 동의 하에 한국전력공사에서 제공하는 ‘한전 파워플래너’ 서비스를 신청하여 확보하였다. 한전 파워플래너는 스마트 전력 계량 인프라(AMI)가 설치된 경우 전력 사용량 및 전기요금 등의 정보를 실시간으로 제공하는 서비스이다. 제품 생산량과 인력 투입시간 및 투입인원 수 데이터는 대상 식품공장에서의 수기로 관리하고 있는 생산일보에서 확보하여 전산화 과정을 거쳐 사용하였으며, 외기온도 및 외기습도 데이터는 기상청에서 제공하는 대상 식품공장이 위치한 지역의 데이터를 사용하였다.

II-2 상관관계 분석

전력 사용량 예측 모델 구현에 앞서 확보한 데이터에 대해 전력 에너지 사용량과의 상관관계를 분석하였다. 변수별 전력 사용량과의 상관관계는 피어슨 상관계수(Pearson Correlation Coefficient)로서 확인하였다. 상관관계수에 대한 해석은 정해진 기준이 없어 데이터 특성, 샘플의 대표성에 따라 연구자의 해석방법이 달라질 수 있다. Rea & Parker[3]는 상관계수가 절댓값을 기준으로 0.0~0.1인 경우 ‘거의 관계가 없음’, 0.1~0.2인 경우 ‘약한 상관관계’, 0.2~0.4인 경우 ‘보통의 상관관계’, 0.4~0.6인 경우 ‘비교적 강한 상관관계’, 0.6~0.8인 경우 ‘강한 상관관계’, 0.8~1.0인 경우 ‘매우 강한 상관관계’를 의미한다고 해석하고 있다. 이에 상관관계가 거의 없다고 볼 수 있는 0.0~0.1 사이의 상관계수를 갖는 데이터를 제외하고, 0.1 이상의 상관관계를 갖는 데이터는 에너지 사용량 예측에 사용하였다.

분석 결과, 전 일(어제) 전력 사용량은 당일의 전력 사용량과 0.62의 상

관계수를 가지며, 전체 제품 생산량은 0.70, 1공장 제품 생산량은 0.70, 2공장 제품 생산량은 0.58, 전체 인력 투입시간은 0.69, 1공장 인력 투입시간은 0.69, 2공장 인력 투입시간은 0.64, 1공장 투입인원 수는 0.67, 2공장 투입인원 수는 0.57, 외기온도는 0.29, 외기습도는 0.12의 상관계수를 갖는 것을 확인하였다. 모든 변수가 모두 당일 전력 사용량과 0.1 이상의 상관계수를 가지며, 이에 전력 사용량 예측에 모든 데이터를 활용되었다. Figure 1에서 각 변수의 전력 사용량에 대한 상관계수를 확인할 수 있다.

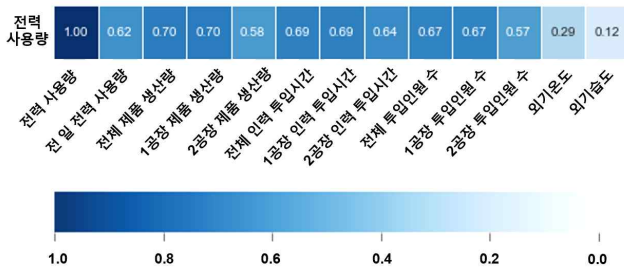


Figure 1. Results of Correlation Analysis

II-3 에너지 소비 예측 모델 구현

에너지 소비 예측 모델은 Linear, Radial Basis Function network(RBF), Polynomial의 3가지 커널로 구분하여 SVR 알고리즘을 적용하였다. Figure 2는 Matplot 라이브러리를 활용하여 SVR 각 커널을 사용했을 시의 데이터 분포를 시각화한 것이다.

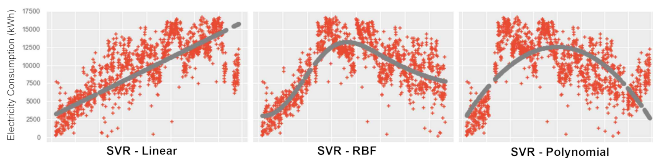


Figure 2. Scatterplots of 3 Case Models (SVR - Linear, RBF, Polynomial)

II-4 에너지 소비 예측 모델 평가

커널별 SVR 알고리즘 기반의 에너지 소비 예측 모델은 미국 냉난방공조 학회(American Society of Heating, Refrigerating and Air-Conditioning Engineers, ASHRAE) Guideline[4]에서 제시하는 평균 제곱근 편차 변동계수(CvRMSE)와 R^2 값을 기준으로 예측 정확도를 평가하였다.

ASHRAE Guideline에서 CvRMSE의 경우, 월 단위(Monthly) 데이터를 사용하는 경우 10%, 시간 단위(Hourly) 데이터를 사용하는 경우 30%를 기준값으로 제시하고 있으며, R^2 의 경우, 0.8을 기준값으로 제시하고 있다. 이에 일 단위(Daily) 데이터를 사용한 본 연구에서는 ASHRAE에서 제시하는 월 단위 및 시간 단위 기준의 중간값인 20%를 CvRMSE 목표값으로 설정하였으며, R^2 는 0.8을 목표값으로 설정하였다.

3가지 케이스에 대해 예측 정확도를 CvRMSE 및 R^2 로 확인한 결과, SVR-Linear는 CvRMSE 21.59%, R^2 0.72, SVR-RBF는, CvRMSE 20.52%, R^2 0.75, SVR-Polynomial는 CvRMSE 22.10%, R^2 0.71이었다 (Figure 3). 이에 세 가지 커널 중 RBF를 적용하였을 때 예측 정확도가 가장 높은 것을 확인하였다.

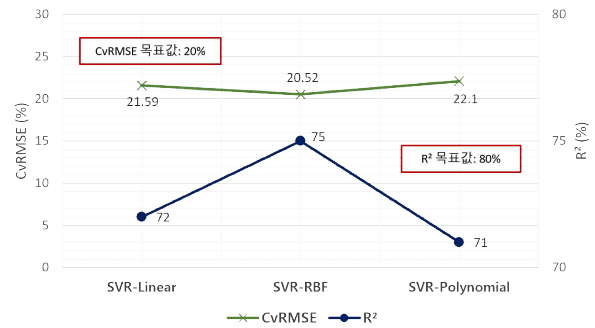


Figure 3. Comparison of Prediction Accuracy of 3 Case Models (SVR - Linear, RBF, Polynomial)

Figure 4는 2017년 1월 1일부터 2017년 1월 31일까지의 한 달에 대해 실제 전력 소비량과 SVR-Linear, SVR-RBF, SVR-Polynomial 모델을 통해 예측된 전력 소비량을 그래프로 나타낸 것으로, 실제값과 SVR 기반 예측 모델을 통한 예측 값이 유사한 그래프 모양을 보이는 것을 확인할 수 있다.

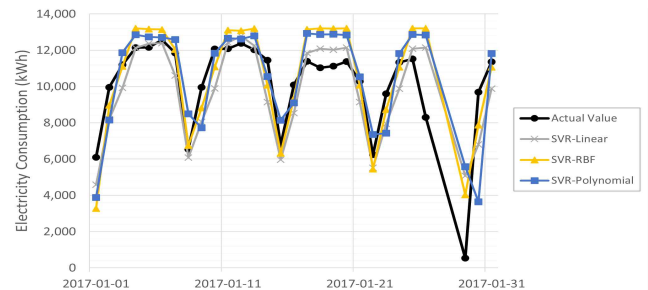


Figure 4. Comparison of Actual Values and Predicted Values

III. 결론

본 논문에서는 한 식품공장을 대상으로, 커널함수에 따른 SVR 알고리즘 기반의 전력 사용량 예측 모델의 성능 비교를 진행하였다. 비교 결과, RBF 커널을 적용하였을 때 CvRMSE 20.52%, R^2 0.75로 예측 정확도가 가장 높은 것을 확인할 수 있었다.

목표값으로 설정한 CvRMSE 20%와 R^2 0.8에 다소 미치지 못하는 결과를 보였으나, 추후 보다 다양한 범주의 데이터를 확보하여 적용하고, 파라미터의 값의 조정을 통해 예측 정확도를 높일 수 있을 것으로 예상된다. 또한 SVR 외의 다양한 알고리즘을 적용하여 대상 식품공장의 에너지 소비 예측을 위한 최적의 모델을 도출할 수 있을 것이다.

ACKNOWLEDGMENT

This work was supported by the Ministry of Trade, Industry and Energy(MOTIE) and the Korea Institute of Energy Technology Evaluation and Planning(KETEP). (No. 20202020800290)

참고 문헌

- [1] 김철우, 김진, 김석민, 황현태, “제조 산업 에너지 절감을 위한 공장 에너지관리시스템(FEMS) 기술 동향 및 적용사례”, 대한설비공학회 설비저널, 44(1), pp.22~27, 2015.
- [2] 산업통상자원부, “제3차 에너지기본계획”, 2019.
- [3] Rea, L., Parker, A., “Designing and Conducting Survey Research : A Comprehensive Guide(3rdEdition)”, John Wiley & Sons, Inc., Jossey-Bass, CA., 2005.
- [4] ASHRAE(American Society of Heating, Refrigerating and Air-Conditioning Engineers), ASHRAE Guideline 14 Measurement of Energy and Demand Savings, pp.4-165, 2002.