

전력사용량 예측 기반 1인 가구 이상 상태 탐지

임혜원, 정상오, 송종협
한국전력공사 데이터사이언스연구소

Detection of Abnormal Status in Single-Person Households based on Electricity Consumption Prediction

Hyewon Lim, Sangoh Jeong, JongHyup Song
Data Science Lab., Korea Electric Power Corporation

Abstract - 취약계층 1인가구의 수가 매년 증가함에 따라 이들에 대한 사회적 안전망 구축에 대한 관심이 높아지고 다양한 기술과 시스템이 개발되고 있다. 전력이나 수도, 가스와 같은 공공재를 기반으로 한 이상탐지는 생활간섭이 없어 사용자의 거부감이 낮고 추가 장비의 투입이 필요없어 서비스의 확대가 용이하다는 장점을 가지고 있다. 본 논문에서는 사용자의 전력사용량을 예측하고 그 결과를 기반으로 이상탐지 조건을 설계하여 사용자의 이상상태를 추정하는 방법을 제시한다. 통계적인 이상탐지 조건의 설계를 위해 통계적 공정관리에서 사용되는 기법들을 적용하였으며, 실험 및 결과를 통해 전력사용량 예측과 이상탐지 조건을 적절히 조화시키기 위한 고려사항들에 대해 살펴보고자 한다.

1. 서 론

우리나라는 2001년 고령인구 비율 7.2%로 고령화 사회에 진입하였고, 2018년에는 그 비율이 14.4%로 고령사회에 들어섰다. 2020년에는 전국 시군구 261개 중 고령인구 비율이 20% 이상인 초고령사회에 진입한 곳이 41.8%로 나타났다. 이러한 고령화 심화에도 불구하고, 부모부양 인식 및 가족 가치관 변화 등으로 독거노인은 증가 추세이며, 만혼이나 비혼과 같은 생활 문화의 변화에 따라 전반적으로 1인가구가 증가하는 추세이다[1]. 2020년에는 1인가구 인구가 617만명으로 전체의 30%를 차지하였으며, 향후 지속적인 증가추세로 예상된다[2].

최근 고독사, 노인자살률 등의 증가에 따라 1인가구에 대한 돌봄 문제가 사회적으로 이슈화되고 있다. 중장년 1인가구의 약 30%는 심각한 사회적 고립이 우려되고 있고, 이들이 결국 노년 1인가구가 될 가능성이 높다. 사회 안전망 구축을 위해 기초지자체에서는 취약 독거노인에 대한 돌봄 서비스를 확대하고 있으나, 돌봄과 지원이 시급한 독거노인의 규모가 빠르게 증가하는 것에 비해 돌봄 서비스 수행인력 부족, 취약 독거노인 수 대비 돌봄 규모 부족 등으로 돌봄 사각지대가 지속적으로 발생하고 있다[1]. 이러한 독거가구의 돌봄 인프라를 강화하기 위해 IoT 기반 서비스를 도입하는 추세이나, 이러한 경우 서비스 운영을 위한 별도의 장비나 센서의 설치가 필요하고, 기기 착용 및 활용을 위한 별도의 사용자 교육이 시행되어야 하거나 프라이버시 문제 등으로 인해 사용자 거부감이 높아질 수 있다는 단점이 있다. 반면, 전기나 수도, 가스와 같이 일상생활의 필수불가결한 요소들을 활용할 경우[3-5] 별도의 기기를 설치할 필요가 없어 돌봄 서비스 유지관리가 용이하고, 사용자에 대한 생활간섭이 없어 사생활 침해에 대한 우려가 없으며 서비스의 확장성이 높다는 장점이 있다.

본 논문에서는 지능형 전력계량 인프라 서비스를 통해 취득된 전력사용량 데이터를 분석하여 사용자의 이상상태를 추정하는 방법을 제시한다. 딥러닝을 기반으로 사용자의 과거 전력사용패턴을 학습하여 미래의 전력사용량을 예측하고, 예측결과를 바탕으로 이상상태 판단을 위한 기준을 설계하였다. 24시간 동안의 해당 기준을 만족하지 못하는 시간이 어느 정도인지 고려하여 사용자의 이상상태를 추정하도록 한다.

2. 본 론

2.1 데이터

2.1.1 데이터 수집

전력사용량 예측 모델을 개발하기 위해 전력 데이터 수집에 동의한 약 80가구의 1시간 단위 부하 프로파일 데이터를 활용하였다. 데이터에는 검침일시, 검침시간, 계기번호, 전력사용량이 포함되며 비식별화 과정을 거쳤다. 통신오류 등으로 인해 발생한 이상치 또는 결측치는 보간법을 사용하여 처리하였다.

2.1.2 학습데이터 구축

전력사용량을 예측하기 위한 학습데이터 구축을 위해 수집한 전력사용량 데이터를 기반으로 시차(Lag) 데이터와 차분(Difference) 데이터를 생성하였다. 시차 데이터는 예측하고자 하는 시간의 직전 시간의 값으로 구성된 데이터이고, 차분 데이터는 직전 시간 값과 그 전 시간 값의 차이값으로 구성된 데이터이다.

본 연구에서는 매 시간 이상을 판단하는 것이 아니라 24시간 단위로 판단을 하는 것을 목표로 하므로, 예측하고자 하는 날짜의 데이터는 모델 학습에 사용할 수가 없다. 따라서 예측하고자 시점을 t 라고 가정할 때, $t-(24 \times n)$ 일, 즉, $t-24$, $t-48$, $t-72$ 등의 데이터를 기반으로 학습데이터를 구성하여야 한다. 본 연구에서 $n=3$ 으로 설정하였다.

2.2 전력사용량 예측모델

전력사용량 예측을 위해서 딥러닝 기반의 예측모델을 생성하고자 하였다. 본 연구에서는 고객별 예측 사용량을 이상탐지를 위한 기준 설계에 사용하고자 하였으므로, 예측성능이 얼마나 좋은지에 초점을 맞추기 보다는, 기준 설계에 얼마나 적합한 결과를 보이는지에 집중하여 최종 모델을 선정하였다. 이를 위해 MLP (Multi Layer Perceptron) 모델과 LSTM (Long Short Term Memory) 모델을 생성한 후, 결과를 비교하는 과정을 거쳤다.

2.2.1 MLP 모델

MLP는 가장 기본적인 형태의 인공신경망 구조로 입력층과 출력층 사이에 하나 이상의 은닉층이 존재하는 신경망이다. 이는 입력값과 목표로 하는 출력값으로 구성된 학습 데이터를 이용하여 가중치를 조정함으로써 목표하는 출력값과 학습 시 출력되는 출력값의 오차가 작아지도록 한다.

본 연구에서는 사용한 MLP 모델은 60개의 뉴런으로 구성된 은닉층 1개로 구성하였으며, 활성화 함수로는 Dying ReLU 문제를 해결하기 위해 LeakyReLU를 사용하였다. 최적화 기법으로는 Adam을 적용하였다.

2.2.2 LSTM 모델

일반적인 순환신경망 모델은 학습 시 최신 정보가 아닌 시간의 격차가 큰 정보를 사용할 경우 역전파시 그라디언트 (gradient)가 점점 줄어들어 학습력이 감소하는 문제가 나타난

다. LSTM은 이러한 문제점들을 해결하기 위하여 긴 기간에 대한 학습이 가능하도록 개선된 신경망 모델이다. 본 연구에서는 LSTM 모델을 2개의 은닉층으로 구성하였으며, 첫 번째 레이어는 16개 노드, 두 번째 레이어는 8개 노드로 구성하였다. 활성화 함수는 MLP와 마찬가지로 LeakyReLU 함수, 최적화 기법은 Adam을 적용하였다.

2.3 이상탐지 조건설계

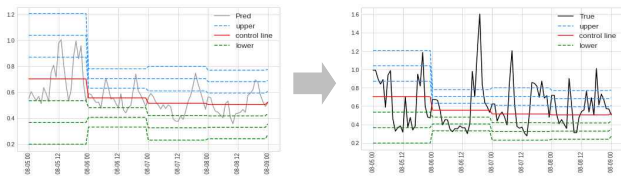
사용자마다 연령, 성별, 환경 등에 따라 생활패턴이 다르기 때문에 1인가구라 하더라도 특정 패턴을 찾아내기 어렵다. 따라서 이상상태를 추정하기 위한 일반적인 조건보다는, 사용자별로 적용가능한 조건으로 설계하여야 한다. 이를 위해서, 본 연구에서는 통계적 공정관리에서 생산 또는 공정의 이상을 판별하기 위한 도구인 관리도(Control Chart)를 활용하여 조건을 설계하였다. 관리도는 시간에 따른 프로세스의 변화를 모니터링하는 차트로 공정 품질 데이터들의 평균선(Control Line), 상한선(Upper Control Limit), 하한선(Lower Control Limit)을 그려 이상 판단의 가이드로 사용한다. 아래의 수식이 이에 해당하며, x 는 공정 품질 값, n 은 데이터 수, a 는 상한선과 하한선의 범위를 지정하기 위한 값이다($a = 1, 2, 3, \dots$).

$$Control\ Line = \frac{1}{n} \sum_{i=1}^n x_i$$

$$Upper/Lower\ Control\ Limit = Control\ Line \pm (a \times \sigma)$$

일반적으로는 데이터가 한계선 밖에 있거나 특별한 패턴을 보일 경우 이상 가능성이 있다고 판단하는데, 이러한 관리도의 개념을 그대로 적용할 경우, 전력사용량이 한계선 밖으로 넘어가면 이상으로 판단되는 문제가 발생하게 된다. 전력사용 패턴의 이상 추정 시에는 일반적인 공정과 달리 사용량 변동이 적거나, 평소보다 전력사용이 적은 경우 등을 탐지하는 것이 목적이기 때문에 이에 맞도록 새로운 조건 설계가 필요하다.

본 연구에서는 전력사용량 예측결과를 기반으로 이상추정을 위한 평균선, 상한선, 하한선을 그리고(그림 1 좌측), 실 사용량을 이와 비교(그림 1 우측)함으로써 이상을 판단할 수 있도록 설계하였다.



〈그림 1〉 이상상태 탐지를 위한 조건 설계

2.4 성능평가

2.4.1 전력사용량 예측 평가

성능평가지표로는 평균절대오차(Mean Absolute Error, MAE), 평균제곱오차(Mean Squared Error, MSE), 평균절대비오차(Mean Absolute Percentage Error, MAPE)를 사용하였다. 성능평가 결과, 표 1에서 볼 수 있듯이 제안한 두 모델 중 MLP가 LSTM에 비해 보다 나은 성능을 보였다. 모델 학습을 위해 사용한 학습 데이터의 복잡도가 높지 않아 간단한 구조의 모델에서도 좋은 성능을 나타내는 것으로 보인다.

〈표 1〉 전력사용량 예측모델 성능평가 결과

	MLP	LSTM
MAE	0.062	0.081
MSE	0.015	0.022
MAPE	24.472	32.987

본 연구에서와 같이 이상판단을 위한 조건을 생성하기 위해 전력사용량을 예측하는 경우에는, 정량적인 성능뿐만 아니라 예측결과가 이상판단 기준 구간을 설정하기에 적절한지를 판단하기 위한 정성적인 성능도 고려해야 한다. 이상상태로 판단되어야 하는 전력사용량의 변화도 잘 예측하는 경우, 예측결과를 바탕으로 이상판단 구간을 설정하여도 실제 사용량이 해당 구간에서 정상으로 판단될 가능성이 높다. 따라서 예측모델 생성 시 전반적인 전력사용 추이를 따라가면서도 변화에 따른 이상 탐지가 가능한 결과를 도출할 수 있도록 할 필요가 있다.

2.4.2 이상상태 추정 성능평가

이상탐지 성능을 평가하기 위해서, 지자체에서 안부를 확인한 결과를 정답으로 활용하였다. 2020년 6월~10월까지 평균 79명을 대상으로 하였으며, 전체 인원을 매일 확인하지 않았고 돌봄대상자별로 이벤트 발생 시 확인하였으므로, 성능평가에 활용할 수 있는 일수는 평균 135.6일(인원수 × 월별 일수 중 안부가 확인된 일수의 평균)이다.

성능평가 결과는 표 2과 같다. 평균 오탐률이 45.6%, 평균 정탐률이 54.4%로 수치상으로 정탐률이 낮다고 볼 수도 있으나, 오경보 중에서 유의미한 경보인 경우가 존재할 수 있어 안부를 살핀다는 의미에서는 오경보임에도 불구하고 경보를 발생시키는 것이 합리적인 경우도 있을 수 있다는 점을 고려해야 한다. 그림 2를 살펴보면, 회색선이 실제 사용량, 녹색선이 예측사용량, 빨간색 네모가 이상으로 탐지된 구간(실제로는 정상상태)이다. 이 구간은 과거에 비해 사용량이 급격히 적어지고 변동량도 적어 이상 알람을 주는 것이 유의미하다고 볼 수 있다.

〈표 2〉 이상탐지 성능평가 결과

	6월	7월	8월	9월	10월	평균
정상탐지	45.2%	60.0%	46.2%	23.1%	53.8%	45.6%
오탐지	54.8%	40.0%	53.8%	76.9%	46.2%	54.4%



〈그림 2〉 이상상태 오탐지 예시

3. 결 론

본 논문에서는 전력사용량을 기반으로 1인가구의 이상상태를 탐지하기 위한 방법을 제안하였다. MLP 기반의 전력사용량 예측 모델을 생성하고, 예측한 전력사용량을 바탕으로 통계적 이상상태 탐지조건을 설계하였다. 성능평가를 통해 정상 이상으로 오탐지한 경우에도 유의미한 결과일 가능성이 있어 일반적인 이상탐지나 시계열 분석과 달리 정량적인 평가 역시 필요함을 보였다. 제안방식은 사용량 변동이 매우 적은 사용자의 이상탐지가 어렵다는 한계가 있어 이에 대한 후속연구가 필요하다.

〈참고 문헌〉

- [1] 보건복지부, “제2차 독거노인 종합지원대책(‘18~’22)”, 2018
- [2] 통계청, “가구주의 연령/가구원수별 추계가구(시도)”, 2019
- [3] 채창훈 외, “독거노인의 전력사용량 실 사례 분석을 통한 임계값 산출 및 무활동 판단 연구”, 2014년도 에너지그리드 정보처리연구회 학술대회 논문지, 제1권, 제1호, pp.106-109, 2014
- [4] 김지현 외, “통합검침데이터 기반 수용가 내 사람의 활동 및 비활동 감지 방법 연구”, 2011년도 대한전기학회 하계학술대회 논문집, pp.1981-1982, 2011
- [5] Gyubaek Kim and Sanghyun Park, “Activity Detection from Electricity Consumption and Communication Usage Data for Monitoring Lonely Deaths,” Sensors, 21(9), 2021