독거노인 돌봄을 위한 전력데이터 기반 기계학습 이상진단 분석

김에덴, 고석갑, 이병탁
한국전자통신연구원

{kimed93, softgear, bytelee}@etri.re.kr

Analysis on machine learning anomaly detection using electricity data for care of elderly living alone

Kim E Den, Ko Seok Kap, Lee Byung Tak
Electronics and Telecommunications Research Institute (ETRI)

요 약

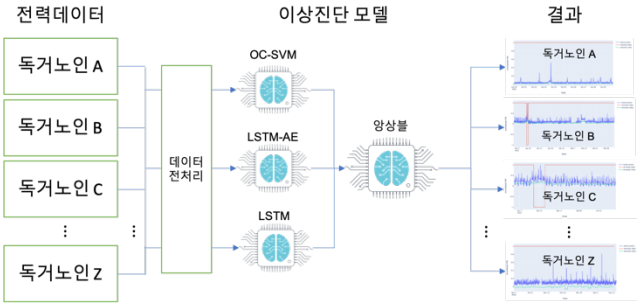
본 논문은 고령화로 인한 독거노인 인구의 증가로 돌봄 서비스에 관한 사회적 관심이 증가함에 따라 독거노인들의 전력데이터의 사용 패턴을 분석하여 딥러닝 및 머신러닝 모델을 통한 이상진단 기술을 소개하고자 한다. 실생활에서 전력 사용은 필수적이고 수집이 용이하기 때문에 전력 사용 패턴을 분석하여 독거노인의 이상상황을 판단한다. 본 논문에서는 세가지의 이상진단 모델을 제시함으로써 각 모델의 결과와 앙상블 기법을 통해 보다 신뢰성 있도록 이상진단 여부를 판별하고 독거노인의 이상 지수를 도출함으로 독거노인의 이상 정도에 대한 정보도 제공한다. 이와 같은 연구를 통하여 추후 사회복지사 또는 독거노인 보호자와의 연계를 통하여 사전 진단 혹은 이상 상황을 감지하여 효율적인 관리 및 서비스 등으로 활용될 수 있다.

Ⅰ. 서론

고령화 시대를 통하여 우리나라의 65 세 이상 인구는 현재 850 만명 정도로 늘어났고 홀로 사는 독거노인의 비율은 19.6%로 계속해서 늘어나고 있는 실정이다[1]. 이처럼 홀로 지내는 노인의 숫자와 비율은 늘어나고 있지만 이를 체계적으로 관리하고 보살 필 수 있는 인프라나 시설은 아직 부족하다. 이와 같은 문제를 해결하기 위해 독거노인 케어 서비스를 위한 기술들을 개발하고 있다. 그 중에 본 논문에서는 독거노인이 생활하는데 필수적인 요소인 전력 데이터를 수집하여 전력 수요 패턴 분석을 통한 AI 이상진단 연구에 대하여 소개한다.

기존의 독거노인 케어는 사람이 직접 확인하거나 전력 혹은 다른 수치들을 직접 모니터링하는 방법으로 대부분 구성되어 있다. 따라서 결국 인적 재원이 가장 중요하게 필요한 요소이다. 이러한 인적 재원을 보완하기 위해 최근 로봇, 스피커 등 많은 기술이 개발되고 있으나 이러한 서비스는 비용이 많이 요구된다는 점이 존재한다. 반면 본 논문의 연구는 전력데이터를 이용하기 때문에 보다 경제적으로 서비스가 가능하며 인공지능을 접목함으로 인적 재원의 손실을 줄일 수 있다는 장점이 있다. 비슷한 사례로 시행되는 서비스도 존재하는데, 본 논문에서는 전력 패턴을 통해 이상진단 결과 외에도 독거노인의 상태를 가늠할 수 있는 지표인 이상 지수를 추출하여 사전 돌봄 관리가 가능하다.

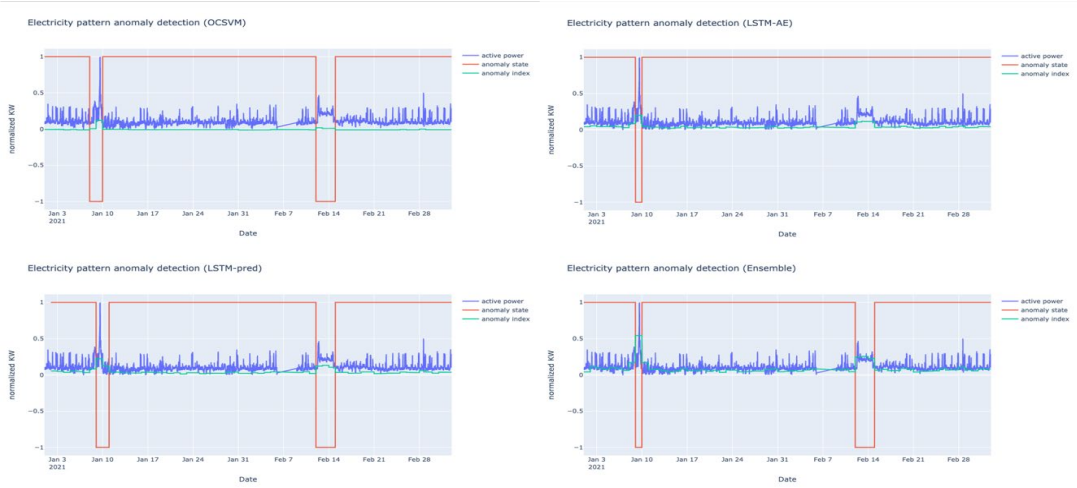
본 논문에서는 전력 패턴 이상진단을 위한 딥러닝 및 머신러닝 모델을 3 가지 사용하였다. 이는 경우에 따라 앙상블하여 단일 모델보다 신뢰성 있는 이상진단 결과를 도출할 수 있다. 이는 하루 단위로 학습되어 하루 전력 사용 패턴을 보고 이상상황을 진단하고 이상 지수를 도출하는데 사용된다.



(그림 1) 이상진단 모델 개념도

Ⅱ. 본론

데이터 및 전처리. 본 연구를 진행하기 위해서는 독거노인 가구에 설치된 AMI 기기를 통해 15 분 단위로 전력량 데이터를 수집한다. 15 분 단위로 수집된 데이터는 그 자체가 전력 소비 패턴을 나타내는 것이 아니기 때문에 하루 단위로 데이터를 재구성하여 분석하는 작업이 필요하다. 즉, 15 분마다 데이터를 수집하므로 하루의 전력 소비 패턴은 96 개의 데이터 포인트를 의미한다. 전력량 데이터를 AMI 기기로 수집할 때 정전 혹은 통신 등의 상황들로 인하여 데이터의 결측치가 발생하기도 한다. 이러한 경우 적절한 값을 채워 넣을 수 있으면 좋지만 장기간 결측치가 발생하는 경우도 있고 임의로 데이터의 값을 채워 넣었을 때, 이상진단의 측면에서는 에러가 높아질 수 있어 부적절하다. 따라서 이러한 경우에는 데이터를 하루단위로 재구성 할 때 하루 단위의 데이터 96 개중에 하나라도 결측치가 있으면 그 하루 전체의 데이터를 사용하지 않는다. 결측치 없이 수집된 정상 데이터만을 필터링하여 이상진단 모델 학습하는데 사용한다.



(그림 2) 데이터셋 2 에 대한 이상진단 및 이상지수 도출 결과

결측치가 제거된 데이터들은 딥러닝 및 머신러닝 모델에 적절히 사용되기 위하여 데이터 정규화 작업이 필요하다. 각 데이터마다 기저 전력과 피크 전력량이 달라 최대-최소 정규화를 통하여 모든 전력 데이터를 0 과 1 사이값으로 정규화를 진행하여 모델 학습 및 테스트 하기 위한 환경을 구축한다.

이상진단 모델. 본 논문에서는 이상진단을 위한 딥러닝 및 머신러닝 모델로 OCSVM 이상진단 방법, LSTM AutoEncoder 를 통한 이상진단 방법, LSTM 예측을 통한 이상진단 방법 이렇게 총 3 가지의 모델을 사용하였다. 먼저 OCSVM 방법은 하나의 클래스로 SVM 기법을 통하여 정상 범주를 정하고 그 범주를 벗어날 때에 이상상황이라고 판단하는 기법이다[2]. LSTM AutoEncoder 이상진단 방법은 LSTM 모델로 구성된 인코더, 디코더를 학습하고 새로운 데이터에 대한 복원 값의 차이를 계산한다. 그 차이의 범주에 따라 정상인지 비정상인지 판별한다. 마지막으로 LSTM 예측을 통한 이상진단 방법은 과거 데이터를 입력으로 LSTM 모델의 결과로 당일 전력패턴을 예측한다. 그 LSTM 예측값이 실제 전력 데이터와 얼마나 차이가 나는지를 계산하고 그 차이의 범주에 따라서 정상인지 비정상인지 판별한다[3]. OCSVM 은 이상진단에 가장 기본 모델로의 역할을 LSTM 은 시계열 특징을 통한 분석으로 오토인코더를 통해 현재 데이터로, 예측모델을 통해 과거 데이터로 이상판단을 함으로 이 세가지의 케이스를 고려하여 이상진단 모델들이 선택되었다.

실험 결과. 본 실험은 6 곳의 독거노인 가구의 전력데이터를 수집하여 진행하였다. 각 데이터 셋은 5~9 개월 동안 수집된 데이터이고 시간으로 트레이닝과 테스트 셋을 나누어서 실험을 진행하였다. 각 데이터의 구성은 다음 표 1 에서 나타낸다.

<표 1> 실험 데이터 정보

	트레이닝 셋	테스트 셋
데이터셋 1	2020.06.16 ~ 2020.09.30	2020.10.01 ~ 2020.10.30
데이터셋 2	2020.06.16 ~ 2020.12.31	2021.01.01 ~ 2021.03.05
데이터셋 3	2020.07.03 ~ 2020.12.31	2021.01.01 ~ 2021.02.06
데이터셋 4	2020.07.03 ~ 2020.12.31	2021.01.01 ~ 2021.03.27
데이터셋 5	2020.07.03 ~ 2020.12.31	2021.01.01 ~ 2021.03.05
데이터셋 6	2020.07.03 ~ 2020.10.31	2020.11.01 ~ 2020.12.15

그림 2 는 대표적으로 데이터셋 2 에 대한 이상진단 결과를 나타낸다. 이는 2020 년 6 월 16 일 부터 12 월 31 일까지의 데이터를 정상으로 학습시킨 후 2021 년 1 월 1 일 부터 3 월 5 일까지의 데이터를 통하여 실험한 결과이다. 각 모델의 결과를 보여주고 최종적으로는 세가지 결과를 통한 앙상블 결과를 나타낸다. 앙상블 모델은 각 모델로부터 추출된 이상지수 결과를 집계한 일종의 배깅 기법을 통하여 최종 결과를 도출하였다.

세부적으로 이상진단 결과를 살펴보면 육안으로도 전력량 패턴이 평소와 다르게 많이 사용된 날에 대부분의 모델은 이상상황이라고 감지하였다. 그림 2 를 통해 보면 1 월 9-10 일, 2 월 12-14 일이 이에 해당한다. LSTM-AE 모델의 경우 2 월 12-14 일을 정상으로 판단하였지만, 이상 지수는 다른 정상 상황에 비해 비교적 높은 값을 나타내고 있어 정상 범주에 있지만 이상상황에 대한 사전대비가 필요함을 알려준다. 최종 앙상블 모델에서는 세가지 모델의 이상지수를 집계하여 1 월 9 일과 2 월 12-14 일의 전력 패턴에 이상이 있다고 감지하였다. 그 중에서도 1 월 9 일의 이상지수가 높기에 이상 정도가 더 높음을 볼 수 있다.

Ⅲ. 결론

본 논문에서는 전력 소비 패턴 데이터를 통하여 독거노인 돌봄을 위한 이상진단 모델을 개발하고 실험하였다. 하루 단위로 데이터를 전처리 하고 3 가지 모델을 통하여 이상진단 및 이상 지수를 추출하여 사전 진단으로도 활용 가능성이 있어 독거노인 돌봄 서비스를 접목할 때 보다 유용한 서비스를 위한 연동이 가능하다. 또한 앙상블 기법을 통하여 단일 모델의 단점을 보완할 수 있어 보다 신뢰성 있는 결과를 도출할 수 있다.

ACKNOWLEDGMENT

본 연구는 산업통상자원부(MOTIE)와 한국에너지기술평가원(KITEP)의 지원을 받아 수행한 연구 과제임. (No. 20181210301570)

참 고 문 헌

[1] 통계청, 「장래가구추계 2017」, 「장래인구추계 2017」 2021.

[2] Li, K., et al, “Improving one-class SVM for anomaly detection.” In Proc. of ICMLC, 2003.

[3] Wang, R., et al, "Deep learning for anomaly detection." In Proc. of WSDM’20, 2020.