(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0093961
(43) 공개일자 2025년06월25일

(51) 국제특허분류(Int. Cl.)
G06N 20/20 (2019.01) *G06N 3/045* (2023.01)
G06Q 50/06 (2024.01)
(52) CPC특허분류
G06N 20/20 (2021.08)
G06N 3/045 (2023.01)
(21) 출원번호 10-2023-0184361
(22) 출원일자 2023년12월18일
심사청구일자 2023년12월18일

(71) 출원인
인천대학교 산학협력단
인천광역시 연수구 갯벌로 27, 인천대학교 이노베이션센터(송도동)
(72) 발명자
김관호
인천광역시 연수구 컨벤시아대로42번길 61, 802동 2402호 (송도동, 송도 더샵 그린애비뉴 8단지)
김종은
인천광역시 부평구 안남로222번길 30, 202동 106호 (산곡동, 경남아파트)
김수희
인천광역시 미추홀구 문화로 51, 101동 203호 (관교동)
(74) 대리인
특허법인충정

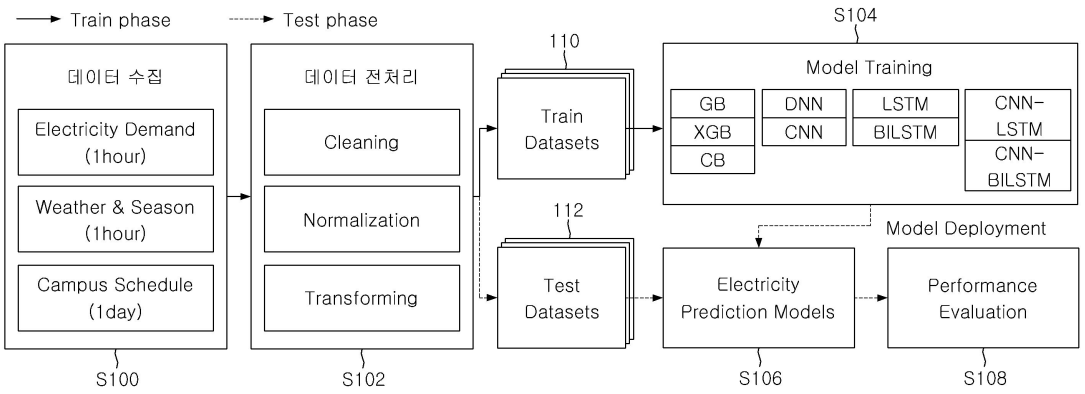
전체 청구항 수 : 총 7 항

(54) 발명의 명칭 **건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법**

(57) 요 약

본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법은, (A) 적어도 하나의 프로세서가, 복수 유형의 건물의 전력 수요 예측을 위한 입력 변수와 관련된 데이터를 수집하는 단계; (B) 상기 수집된 데이터를 전처리하는 단계; (C) 상기 전처리된 데이터를 훈련 데이터 세트로 사용하여 복수의 기계학습 및 딥러닝 알고리즘에 기반한 전력 수요 예측 모델들을 학습시키는 단계; (D) 학습 과정에 사용되지 않은 전처리된 데이터들을 테스트 데이터 세트로 사용하여, 학습된 복수의 전력 수요 예측 모델 각각이 전력 수요를 예측하는 단계; 및 (E) 복수의 전력 수요 예측 모델 각각이 예측한 전력 수요와 실제 전력 사용량에 기반하여 복수의 전력 수요 예측 모델 각각의 예측 성능을 평가하는 단계를 포함할 수 있다.

대 표 도



(52) CPC특허분류
G06Q 50/06 (2024.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1415186773
과제번호 20212020900090
부처명 산업통상자원부
과제관리(전문)기관명 한국산업기술평가관리원
연구사업명 에너지수요관리핵심기술개발(에특)
연구과제명 에너지 다소비 건물의 에너지 관리 및 최적화를 지원하는 국내외 기술표준 연계형 디지털 트윈 플랫폼 및 사업 모델 개발
기 여 율 1/1
과제수행기관명 (주)에코시안
연구기간 2023.01.01 ~ 2023.12.31

명 세 서

청구범위

청구항 1

(A) 적어도 하나의 프로세서가, 복수 유형의 건물의 전력 수요 예측을 위한 입력 변수와 관련된 데이터를 수집하는 단계;

(B) 상기 수집된 데이터를 전처리하는 단계;

(C) 상기 전처리된 데이터를 훈련 데이터 세트로 사용하여 복수의 기계학습 및 딥러닝 알고리즘에 기반한 전력 수요 예측 모델들을 학습시키는 단계;

(D) 학습 과정에 사용되지 않은 전처리된 데이터들을 테스트 데이터 세트로 사용하여, 학습된 복수의 전력 수요 예측 모델 각각이 전력 수요를 예측하는 단계; 및

(E) 복수의 전력 수요 예측 모델 각각이 예측한 전력 수요와 실제 전력 사용량에 기반하여 복수의 전력 수요 예측 모델 각각의 예측 성능을 평가하는 단계를 포함하는, 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법.

청구항 2

제1항에 있어서,

상기 건물의 전력 수요 예측을 위한 입력 변수는 기후 및 계절 정보, 건물 운영 정보, 피크 시간 이전의 전력 사용량을 포함하는, 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법.

청구항 3

제2항에 있어서,

상기 기후 및 계절 정보는 외기 온도, 습도, 풍속, 일사량, 전운량 및 계절 정보를 포함하고,

상기 건물 운영 정보는 학기, 방학, 요일에 따른 건물의 재실률을 고려하기 위해 요일 정보와 학기 운영 정보를 포함하는, 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법.

청구항 4

제1항에 있어서,

상기 복수의 전력 수요 예측 모델의 출력 변수는 입력 변수와 같은 날의 피크 시간에 예측된 전력 사용량을 포함하는, 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법.

청구항 5

제1항에 있어서,

상기 복수의 기계학습 및 딥러닝 알고리즘은,

그래디언트 부스팅 기법, XGBoost 기법, CatBoost 기법, 심층신경망(DNN), 합성곱 신경망(CNN), 장단기 기억 신경망(LSTM), 양방향 장단기 기억 신경망(BILSTM), CNN-LSTM 및 CNN-BILSTM을 포함하는, 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법.

청구항 6

제1항에 있어서,

상기 복수 유형의 건물은, 전력 사용량 패턴의 차이가 존재하는 세 가지 유형의 건물을 포함하고,

상기 세 가지 유형의 건물은, 강의 스케줄에 따라 유동 인구의 변동성이 커 가장 높은 변동성을 보이는 강의실로 구성된 건물, 낮은 전력 사용량을 보이며 변동성이 작은 연구실로 구성된 건물, 및 평균 전력 사용량이 가장

높지만 강의실에 비해 낮은 변동성을 보이는 행정 관리실을 포함하는, 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법.

청구항 7

제6항에 있어서,

BILSTM을 사용하여, 평균 전력 사용량이 가장 높지만 강의실에 비해 낮은 변동성을 보이는 행정 관리실의 전력 수요를 예측하고,

CNN을 사용하여, 낮은 전력 사용량을 보이며 변동성이 작은 연구실의 전력 수요를 예측하고,

BILSTM을 사용하여, 강의 스케줄에 따라 유동 인구의 변동성이 커 가장 높은 변동성을 보이는 강의실의 전력 수요를 예측하는, 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법에 관한 것이다.

배 경 기 술

[0002] 한국에너지공단에 따르면 대학 건물의 전력 사용량은 코로나 19 시기를 제외하고 매년 증가하고 있으며 대학은 건물 업종별 에너지 사용량 중 세 번째로 높은 비중을 차지하는 에너지 다소비 건물로 고려된다. 대학 건물의 전력 사용량은 2022년 전년 대비 7.7%의 높은 상승률을 기록하였으며 이는 교육용 전력 사용량의 43%에 해당한다. 더 나아가 국외 대학의 건물 역시 높은 전력 사용량을 보이며 대학 캠퍼스의 에너지 관리 및 절감이 요구되고 있다. 정확한 대학 건물의 전력 수요 예측은 건물의 에너지 관리와 안정적인 에너지 공급을 위해 필수적이다. 최근 대학들은 건물의 에너지 관리를 위해 건물에너지관리시스템(Building Energy Management System, BEMS)을 도입하는 추세이다.

[0003] BEMS는 에너지 사용량을 분석하여 전력 수요를 예측하고 이를 통해 효율적으로 에너지를 관리하는 통합시스템으로 건물의 전력 수요 예측은 BEMS의 효과적인 운영을 위한 필수 구성 요소이다. 전력 사용량 예측을 통해 미래 수요를 고려한 에너지 운영 계획을 수립할 수 있으며 수요에 맞는 공급을 통해 에너지 낭비와 생산 비용을 최소화할 수 있다. 그러나 각 건물의 전력 수요는 건물의 특성과 기후 조건에 따라 수요 패턴의 차이가 존재하기

[0004] 때문에 건물의 전력 수요 예측은 매우 어려운 문제로 고려된다. 예를 들어, 유동 인구가 많은 학기 중에는 전력 사용량이 많고 변동성이 높지만, 유동 인구가 상대적으로 적은 방학에는 학기 중과 다른 수요 형태를 보인다. 이러한 전력 수요 패턴의 차이는 예측 모델의 정확도에 영향을 미친다. 따라서 건물의 전력 사용량 수요 예측 정확도를 향상시키기 위해서는 수요 패턴별 예측 모델의 성능 비교 연구가 필요하다.

[0005] 또한, 기존 연구들은 새로운 기법을 활용한 건물의 전력 사용량 예측 모델을 통해 예측 성능을 향상시키거나 단일 건물을 대상으로 여러 예측 모델들의 성능을 비교하는 연구가 주로 선행되었다. 단일 건물을 대상으로 수행된 연구는 특정 건물에만 편향된 평가로 이어질 수 있으며 다양한 전력 수요 패턴에 있어 예측 성능이 보장되지 않는다. 또한, 수요 패턴에 따라 모델의 예측 정확도의 차이가 존재하기 때문에 건물의 전력 사용량 패턴에 따른 적합한 모델 선정이 필요하다. 이러한 이유로 기존 연구되었던 예측 모델을 실제 건물에 재현시 제한된 성능을 보임에 따라, 건물의 수요 패턴에 따른 전력 사용량 예측 모델의 성능 비교 연구가 요구된다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) KR 10-1168153 B1

발명의 내용

해결하려는 과제

[0007] 본 발명이 해결하고자 하는 과제는 건물의 전력 수요 패턴에 따른 예측 성능을 비교하기 위해 다양한 전력 수요 패턴을 보이는 캠퍼스 건물들을 대상으로 기계 학습 및 딥러닝 기반의 전력 수요예측 모델의 예측 성능을 비교할 수 있는 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법을 제공하는 것이다.

과제의 해결 수단

[0008] 상기 과제를 해결하기 위한 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법은,

[0009] (A) 적어도 하나의 프로세서가, 복수 유형의 건물의 전력 수요 예측을 위한 입력 변수와 관련된 데이터를 수집하는 단계;

[0010] (B) 상기 수집된 데이터를 전처리하는 단계;

[0011] (C) 상기 전처리된 데이터를 훈련 데이터 세트로 사용하여 복수의 기계학습 및 딥러닝 알고리즘에 기반한 전력 수요 예측 모델들을 학습시키는 단계;

[0012] (D) 학습 과정에 사용되지 않은 전처리된 데이터들을 테스트 데이터 세트로 사용하여, 학습된 복수의 전력 수요 예측 모델 각각이 전력 수요를 예측하는 단계; 및

[0013] (E) 복수의 전력 수요 예측 모델 각각이 예측한 전력 수요와 실제 전력 사용량에 기반하여 복수의 전력 수요 예측 모델 각각의 예측 성능을 평가하는 단계를 포함할 수 있다.

[0014] 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법에 있어서,

[0015] 상기 건물의 전력 수요 예측을 위한 입력 변수는 기후 및 계절 정보, 건물 운영 정보, 피크시간 이전의 전력 사용량을 포함할 수 있다.

[0016] 또한, 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법에 있어서,

[0017] 상기 기후 및 계절 정보는 외기 온도, 습도, 풍속, 일사량, 전운량 및 계절 정보를 포함하고,

[0018] 상기 건물 운영 정보는 학기, 방학, 요일에 따른 건물의 재실률을 고려하기 위해 요일 정보와 학기 운영 정보를 포함할 수 있다.

[0019] 또한, 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법에 있어서,

[0020] 상기 복수의 전력 수요 예측 모델의 출력 변수는 입력 변수와 같은 날의 피크 시간에 예측된 전력 사용량을 포함할 수 있다.

[0021] 또한, 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법에 있어서, 상기 복수의 기계학습 및 딥러닝 알고리즘은,

[0022] 그래디언트 부스팅 기법, XGBoost 기법, CatBoost 기법, 심층신경망(DNN), 합성곱 신경망(CNN), 장단기 기억 신경망(LSTM), 양방향 장단기 기억 신경망(BILSTM), CNN-LSTM 및 CNN-BILSTM을 포함할 수 있다.

[0023] 또한, 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법에 있어서,

[0024] 상기 복수 유형의 건물은, 전력 사용량 패턴의 차이가 존재하는 세 가지 유형의 건물을 포함하고,

[0025] 상기 세 가지 유형의 건물은, 강의 스케줄에 따라 유동 인구의 변동성이 커 가장 높은 변동성을 보이는 강의실로 구성된 건물, 낮은 전력 사용량을 보이며 변동성이 작은 연구실로 구성된 건물, 및 평균 전력 사용량이 가장 높지만 강의실에 비해 낮은 변동성을 보이는 행정 관리실을 포함할 수 있다.

[0026] 또한, 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법에 있어서,

[0027] BILSTM을 사용하여, 평균 전력 사용량이 가장 높지만 강의실에 비해 낮은 변동성을 보이는 행정 관리실의 전력

수요를 예측하고,

[0028] CNN을 사용하여, 낮은 전력 사용량을 보이며 변동성이 작은 연구실의 전력 수요를 예측하고,

[0029] BILSTM을 사용하여, 강의 스케줄에 따라 유동 인구의 변동성이 커 가장 높은 변동성을 보이는 강의실의 전력 수요를 예측할 수 있다.

발명의 효과

[0030] 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법에 의하면, 다양한 전력 수요 패턴을 보이는 캠퍼스 건물들을 대상으로 기계 학습 및 딥러닝 기반의 전력 수요예측 모델의 예측 성능을 비교함으로써, 다양한 전력 수요 패턴을 보이는 건물 유형별로 최적의 전력 수요 예측 모델을 결정할 수 있다.

[0031] 또한, 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법에 의하면, 다양한 전력 수요 패턴을 보이는 건물 유형에 따라 최적의 전력 수요 예측 모델을 사용함으로써, 건물 유형에 따라 전력 수요를 정확하게 예측할 수 있다.

도면의 간단한 설명

[0032] 도 1은 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법의 개요도.

도 2는 DNN의 구조를 도시한 도면.

도 3은 CNN의 구조를 도시한 도면.

도 4는 LSTM의 구조를 도시한 도면.

도 5는 BILSTM의 구조를 도시한 도면.

도 6은 한 달 동안 각 건물의 전기 소비 패턴을 도시한 그래프.

도 7은 테스트 데이터에 대한 예측 모델들의 예측 결과의 예를 도시한 그래프.

도 8은 계절에 따른 예측 모델들의 MAE에 대한 박스 플롯을 도시한 도면.

도 9는 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법이 수행될 수 있는 장치를 도시한 도면.

발명을 실시하기 위한 구체적인 내용

[0033] 본 발명의 목적, 특정한 장점들 및 신규한 특징들은 첨부된 도면들과 연관되어지는 이하의 상세한 설명과 바람직한 실시예들로부터 더욱 명백해질 것이다.

[0034] 이에 앞서 본 명세서 및 청구범위에 사용된 용어나 단어는 통상적이고 사전적인 의미로 해석되어서는 아니되며, 발명자가 그 자신의 발명을 가장 최선의 방법으로 설명하기 위해 용어의 개념을 적절하게 정의할 수 있는 원칙에 입각하여 본 발명의 기술적 사상에 부합되는 의미와 개념으로 해석되어야 한다.

[0035] 본 명세서에서 각 도면의 구성요소들에 참조번호를 부가함에 있어서, 동일한 구성 요소들에 한해서는 비록 다른 도면상에 표시되더라도 가능한 한 동일한 번호를 가지도록 하고 있음에 유의하여야 한다.

[0036] 또한, "제1", "제2", "일면", "타면" 등의 용어는, 하나의 구성요소를 다른 구성요소로부터 구별하기 위해 사용되는 것으로, 구성요소가 상기 용어들에 의해 제한되는 것은 아니다.

[0037] 또한, 본 명세서에 개시된 실시예를 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 명세서에 개시된 실시예의 요지를 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다. 또한, 첨부된 도면은 본 명세서에 개시된 실시예를 쉽게 이해할 수 있도록 하기 위한 것일 뿐, 첨부된 도면에 의해 본 명세서에 개시된 기술적 사상이 제한되지 않으며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0038] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시형태를 상세히 설명하기로 한다.

[0039] 본 발명은 건물의 전력 수요 패턴에 따른 예측 성능을 비교하기 위해 다양한 전력 수요 패턴을 보이는 캠퍼스 건물들을 대상으로 기계 학습 및 딥러닝 기반의 전력 수요예측 모델의 예측 성능을 비교하였다. 통계 기법을 활용한 모델에 비해 기계 학습 모델의 성능이 높았던 선행 연구들이 존재하기 때문에 통계적 기법은 제외하였으며 최근 전력 수요 예측 분야에 적용되는 기계 학습 모델(Gradient Boosting, XGBoost, CATBoost)과 딥러닝(DNN, CNN, LSTM, BiLSTM, CNN-LSTM, CNN-BiLSTM) 기반의 전력 수요예측 모델 9개의 예측 성능을 비교하였다.

[0040] 전력 수요 예측 모델은 전력 사용량이 집중되는 피크 시간의 전력 수요를 예측하도록 설계하였다. 피크 시간의 전력 수요는 정전과 에너지 손실이 빈번히 발생하는 구간으로 에너지 관리를 위한 중요한 요소로 고려된다. 예측 모델의 정확도는 회귀 문제에서 자주 사용되는 평가지표 MAE(mean absolute error), RMSE(root mean squared error), MAPE(mean absolute percentage error)를 통해 평가하고 건물별 예측 성능을 비교하였다.

[0041] 성능 비교에 앞서 건물의 전력 수요 예측 모델을 위해 활용된 기계 학습 및 딥러닝 모델에 대해 설명한다. 우선, 전력 수요 예측 모델의 입력 변수와 출력 변수를 설명하고, 예측 성능 비교를 위한 프레임워크에 대해 다룬다. 마지막으로, 예측 모델로 활용된 부스팅 기반의 기계 학습과 딥러닝 모델에 대해 간략히 설명한다.

[0042] 건물의 전력 수요 예측을 위해 고려된 입력 변수와 출력 변수는 표 1과 같다.

표 1

Table 1. Inputs and output variables of building's electricity prediction models.

	Categories	Variables	Notations
Inputs	Weather and Seasonal Information	Temperature	$x_{i,d}$
		Humidity	
		Wind speed	
		Insolation	
		Cloudiness Index	
		Season Index	
	Building Operation Information	Weekday Index	
		Vacation Index	
	Electricity Demand before Peak Time	Electricity demand	
Outputs	Electricity Demand in Peak Time	Predicted Electricity demand	$\hat{y}_{j,d}$

[0043]

[0044] 입력 변수(Inputs)는 기후 및 계절 정보, 건물 운영 정보, 피크 시간 이전의 전력 사용량으로 구성된다. 기후 정보는 전력 사용량에 영향을 미치는 주요 요인으로 고려됨에 따라, 외기 온도, 습도, 풍속, 일사량, 전운량, 계절 정보를 포함하였다. 학기, 방학, 요일에 따른 건물의 재실률을 고려하기 위해 요일 정보와 학기 운영 정보를 포함하였으며 예측 당일 전력 사용량 패턴을 가장 직접적으로 보이는 피크 시간 이전의 전력 사용량을 입력 변수로 고려하였다. 이때, i와 d는 각각 피크 시간 이전 시간과 예측 날짜를 의미하며 $x_{i,d}$는 d일의 i시간에 해당하는 입력 변수를 나타낸다.

[0045] 건물의 전력 수요예측 모델의 출력 변수(Outputs)는 피크 시간에 해당하는 전력 사용량으로 $\hat{y}_{j,d}$는 입력 변수와 같은 날의 피크 시간 j의 예측된 전력 사용량을 의미한다. 전력 피크 시간은 하루 중 전력 수요가 급증하는 시간으로 피크 시간에 대한 여러 정의가 존재하나 본 발명에서는 수집한 데이터를 기반으로 피크 시간을 정의하였다.

[0046] 도 1은 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법의 개요도이다.

[0047] 한편, 도 9는 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법이 수행될 수 있는 장치를 도시한 도면이다.

[0048] 도 9를 참조하면, 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법이 수행될 수 있는 장치(900)는, 하드웨어, 소프트웨어, 또는 이들의 결합으로 이루어질 수 있다. 예를 들어, 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법이 수행될 수 있는 장치(900)는, 위와 같은 기능/단계/과정들을 수행하기 위한 적어도 하나의 프로세서를 포함하는 컴퓨터 또는 인터넷상의 서버 형태로 구현될 수 있다.

[0049] 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법이 수행

될 수 있는 장치(900)는, 버스(910)를 통해 연결되는 적어도 하나의 프로세서(902), 메모리(904), 입출력 인터페이스(908), 및 네트워크 통신 인터페이스(906)를 포함할 수 있다.

[0050] 상기 메모리(904)는 상기 프로세서(902)에 의해 수행될 때 상기 장치(900)가 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법이 수행될 수 있는 장치를 수행하기 위한 특정 동작을 구현하도록 구성된 명령어를 포함하고, 상기 특정 동작은, 복수 유형의 건물의 전력 수요 예측을 위한 입력 변수와 관련된 데이터를 수집하는 단계, 상기 수집된 데이터를 전처리하는 단계, 상기 전처리된 데이터를 훈련 데이터 세트로 사용하여 복수의 기계학습 및 딥러닝 알고리즘에 기반한 전력 수요 예측 모델을 학습시키는 단계, 학습 과정에 사용되지 않은 전처리된 데이터들을 테스트 데이터 세트로 사용하여, 학습된 복수의 전력 수요 예측 모델 각각이 전력 수요를 예측하는 단계 및 복수의 전력 수요 예측 모델 각각이 예측한 전력 수요와 실제 전력 사용량에 기반하여 복수의 전력 수요 예측 모델 각각의 예측 성능을 평가하는 단계를 수행하는 것을 포함할 수 있다.

[0051] 프로세서(902)는 메모리(904)에 저장된 명령어들에 대한 처리를 실행하는 반도체 장치일 수 있다. 메모리(904)는 다양한 종류의 휘발성 또는 불휘발성 저장 매체를 포함할 수 있다. 예를 들어, 메모리(404)는 ROM(Read Only Memory), RAM(Random Access Memory) 또는 자기 디스크 저장 장치와 플래시 메모리 장치와 같은 데이터 저장 장치 등을 포함할 수 있다. 또한, 메모리(904)는 프로세서(902)로부터 떨어져 위치하는 저장장치나, 인터넷 등의 통신 네트워크를 통하여 액세스되는 네트워크 부착형 저장장치 등을 더 포함할 수 있다.

[0052] 또한, 네트워크 통신 인터페이스(906)는 스마트폰, 노트북 PC, 데스크탑 PC 등 사용자 단말에서의 유선 인터넷 통신이나 WiFi, WiBro 등 무선 인터넷 통신, WCDMA, LTE, 5G 등 이동통신을 지원하는 모뎀 등의 통신 모듈이나, 근거리 무선 통신 방식(예, 블루투스, 지그비, 와이파이 등)의 통신을 지원하는 모뎀 등의 통신모듈을 포함할 수 있다.

[0053] 따라서, 본 명세서에 개시된 실시예들과 관련하여 설명된 방법 또는 알고리즘의 단계는 프로세서(902)에 의해 실행되는 하드웨어, 소프트웨어 모듈, 또는 그 2개의 결합으로 직접 구현될 수 있다. 소프트웨어 모듈은 RAM 메모리, 플래시 메모리, ROM 메모리, EPROM 메모리, EEPROM 메모리, 레지스터, 하드 디스크, 착탈형 디스크, CD-ROM과 같이 컴퓨터 등 장치로 판독 가능한 저장/기록 매체(즉, 메모리(904))에 상주할 수도 있다. 예시적인 메모리(904)는 프로세서(902)에 커플링되며, 그 프로세서(902)는 메모리(904)부터 정보(코드)를 판독할 수 있고 메모리(904)에 정보(코드)를 기입할 수 있다. 다른 방법으로, 메모리(904)는 프로세서(902)와 일체형일 수도 있다. 프로세서(902) 및 메모리(904)는 주문형 집적회로(ASIC) 내에 상주할 수도 있다. ASIC는 사용자 단말기 내에 상주할 수도 있다. 다른 방법으로, 프로세서(902) 및 메모리(904)는 사용자 단말기 내에 개별 컴포넌트로서 상주할 수도 있다.

[0054] 입출력 인터페이스(908)는 입력 및 출력 기능을 제공하는 입출력 장치에 연결될 수 있고, 네트워크 통신 인터페이스(906)는 네트워크에 연결된 기기들과 통신할 수 있는 기능을 제공할 수 있다.

[0055] 다시, 도 1을 참조하면, 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법은 크게 학습 과정(Train phase)과 테스트 과정(Test phase)으로 구성될 수 있다.

[0056] S100 단계에서, 적어도 하나의 프로세서(902)가, 복수 유형의 건물의 전력 수요 예측을 위한 입력 변수와 관련된 데이터를 수집할 수 있다.

[0057] S102 단계에서, 상기 수집된 데이터를 전처리할 수 있다.

[0058] 학습 과정의 첫 번째로 수집된 데이터는 모델의 입력으로 활용될 수 있도록 데이터 전처리 과정을 거친다. 전처리 과정에서는 이상치 데이터 제거를 위한 클리닝(114), 데이터의 규모를 균일하게 변환하기 위한 최소-최대 정규화(116) 및 입력 변수 형태로 전환하는 데이터 변환 작업(118)이 수행될 수 있다.

[0059] S104 단계에서, 상기 전처리된 데이터를 훈련 데이터 세트(110)로 사용하여 복수의 기계학습 및 딥러닝 알고리즘에 기반한 전력 수요 예측 모델들을 학습시킬 수 있다.

[0060] 상기 복수의 기계학습 및 딥러닝 알고리즘은, 그래디언트 부스팅 기법, XGBoost 기법, CatBoost 기법, 심층신경망(DNN), 합성곱 신경망(CNN), 장단기 기억 신경망(LSTM), 양방향 장단기 기억 신경망(BILSTM), CNN-LSTM 및 CNN-BILSTM을 포함할 수 있다.

[0061] 전처리된 데이터 중 훈련 데이터(110)는 전력 수요 예측 모델의 학습에 사용되며 예측 모델들은 피크 시간의 전

력 사용량 예측값과 실제 값의 차이를 최소화하는 방향으로 학습할 수 있다.

[0062] 테스트 단계(Test phase)에서는 학습 과정에 사용되지 않은 데이터들을 통해 학습된 예측 모델들의 성능을 평가할 수 있다. 학습 과정과 동일하게 테스트 데이터에 대해 전처리 과정이 수행될 수 있다.

[0063] S106 단계에서, 학습 과정에 사용되지 않은 전처리된 데이터들을 테스트 데이터 세트(112)로 사용하여, 학습된 복수의 전력 수요 예측 모델 각각이 전력 수요를 예측할 수 있다.

[0064] S108 단계에서, 복수의 전력 수요 예측 모델 각각이 예측한 전력 수요와 실제 전력 사용량에 기반하여 복수의 전력 수요 예측 모델 각각의 예측 성능을 평가할 수 있다. 이 단계에서는, 테스트 데이터(112)에 대한 예측 결과를 다양한 실험 세팅을 통해 각 모델의 예측 성능을 비교 분석할 수 있다.

[0065] 하기에, 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법에서 사용될 수 있는 전력 사용량 예측을 위한 기계학습 및 딥러닝 알고리즘에 대해 설명하기로 한다.

[0066] 부스팅 기법은 여러 약한 학습기들을 결합하여 성능이 강한 학습기를 만드는 알고리즘으로 이동 시간, 고객 수요, 전력 수요 등 예측 문제에서 많이 활용된다. 그래디언트 부스팅(GB: Gradient Boosting)은 부스팅 기법과 경사 하강법을 결합한 모델로 이전 예측기의 실제값과 예측값의 차이인 잔여 오차를 학습하는 약한 학습기를 반복 결합하여 미분 가능한 손실함수의 기울기를 최소화하도록 학습한다. 최종 목푯값은 손실함수의 값을 최소화하는 전력 사용량 예측 근사치 $\hat{F}^{*}(x)$ 이며 수학식 1과 같이 계산될 수 있다.

수학식 1

[0067] $$F_m^*(x) = F_{m-1}^*(x) + p_m h_m(x)$$

[0068] F_{m-1}^* 은 m-1번째 예측 근사치를 의미하고, p_m 은 m번째 손실함수의 가중치, h_m 은 손실함수를 최소화하는 함수이다.

[0069] XGBoost(XGB)는 목적함수에 정규화 매개변수를 두어 기존 부스팅의 과적합 문제를 완화한 부스팅 기법이다. 균형 트리 분할 방식을 사용하여 안정적인 학습이 가능하고 병렬처리를 통한 기존 부스팅 기법보다 빠른 학습이 가능하다. 정규화 매개변수를 포함한 XGBoost의 손실함수는 수학식 2와 같다.

수학식 2

[0070] $$L_{xgb} = \sum_{i=1}^{N} L(y_i, F(x_i)) + \sum_{m=1}^{M} \Omega(h_m),$$

$$where\, \Omega(h) = \gamma T + \frac{1}{2}\lambda \parallel w \parallel^2$$

[0071] $L(y_i, F(x_i))$ 는 예측값과 실제값의 차이를 나타내는 손실함수를 의미한다. Ω 는 복잡한 모델을 제어하는 정규화 함수, T는 트리 구조의 잎 노드의 수, λ 는 L2 정규화 파라미터, w 는 잎 노드의 가중치, γ 는 최소 손실 감소를 의미한다.

[0072] CatBoost(CB)는 과적합 방지와 범주형 데이터 처리에 초점을 둔 부스팅 기법이다. 모든 학습 데이터를 대상으로 잔여 오차를 계산하는 기존 부스팅 기법과 달리 CatBoost는 학습 데이터의 일부만으로 잔여 오차를 계산하고 모델을 생성함으로써 과적합을 방지한다. 또한, 데이터 전처리 단계에서 수행되는 범주형 데이터 처리를 학습 과정에서 자동으로 수치형 데이터로 변환하여 데이터 전처리 시간을 단축하였다.

[0073] 심층신경망(DNN)은 입력층과 출력층 사이에 2개 이상의 은닉층으로 구성된 인공신경망이다. 심층신경망은 여러 은닉층을 통해 복잡한 패턴을 학습한다. 도 2와 같이 심층신경망은 여러 층으로 구성되며 각 층은 여러 개의 노

드로 구성된다.

[0074] 각 노드는 이전 층으로부터 입력 변수를 받아 가중치와 활성화 함수를 통해 출력값을 계산하고 생성된 출력값은 다음 층의 입력값으로 사용된다.

[0075] 각 노드의 출력값 u_p는 수학식 3과 같이 계산된다.

수학식 3

[0076] $$u_p = A\left(\sum_{p=1}^{P}(w_{q,p}o_p)\right)$$

[0077] $w_{p,q}$는 노드 p와 q 사이의 가중치를 나타내고, o_p는 노드 p의 출력값을 의미한다. P는 노드 q와 연결된 노드의 수를 나타내고, A는 입력과 출력 간의 비선형 변환을 가능하게 하는 활성화 함수를 의미한다.

[0078] 노드의 가중치는 초기 0과 1 사이의 임의의 값으로 무작위 할당되고 오차 역전파 알고리즘을 활용한 학습 과정을 통해 손실함수의 기울기를 줄이도록 갱신된다. 손실함수는 회귀 문제에서 평균 제곱 오차가 가장 널리 사용되며 그 외에도 문제 상황에 따라 크로스 엔트로피, 휴버, 힌지 등 여러 손실함수가 사용된다. 노드와 은닉층의 개수는 심층신경망의 성능에 큰 영향을 미치는 요소로 문제의 복잡성을 고려한 반복 실험을 통해 결정된다.

[0079] 합성곱 신경망(CNN)은 이미지 데이터를 정형 데이터로 변환할 때 발생하는 데이터 특성 손실을 방지하기 위해 개발된 신경망이다. 합성곱 신경망은 합성곱층과 풀링층을 통해 데이터 내 잠재적 특징을 추출하고 패턴을 학습함으로써 이미지 분류, 인식, 자연어 처리 등 여러 분야에서 활용되고 있다. 최근 합성곱 신경망이 시계열 예측 분야에서 좋은 성능을 보임에 따라 합성곱 신경망을 활용한 건물 에너지 예측 모델 역시 많이 연구되고 있다.

[0080] 합성곱 신경망 구조는 도 3과 같이 합성곱층, 풀링층, 완전 연결 층과 입력층, 출력층으로 구성된다. 합성곱층은 입력 변수들을 대상으로 커널이라고 부르는 가중치 행렬을 이동시키며 합성곱을 수행한다. 합성곱 연산은 수학식 4와 같이 수행된다. 합성곱 연산 결과는 비선형성을 부여하는 활성화 함수를 거쳐 입력 데이터의 특징과 패턴의 활성화 정도를 나타내는 특성맵을 생성한다.

수학식 4

[0081] $$y(k) = \sum_{i=0}^{N} h(k-i)u(i)$$

[0082] h는 커널과 매칭되는 입력 변수들을 의미하고, k는 h에 해당하는 입력 변수의 인덱스를 의미한다. $u(i)$, N은 각각 커널과 커널의 길이를 나타내고, y는 k번째 합성곱 연산 결과를 의미한다.

[0083] 풀링층은 합성곱층에 의해 추출된 특성맵을 입력 받아 크기를 줄이는 다운샘플링을 수행한다. 풀링 연산은 일반적으로 정해진 크기 안에서 가장 큰 값을 추출하는 최대 풀링과 평균값을 추출하는 평균 추출이 사용된다. 풀링 연산을 통해 특징 맵의 중요한 특징을 보존함과 동시에 계산 효율성을 높일 수 있다. 풀링층의 출력값은 완전 연결층의 입력 변수로써 사용되며 완전 연결층은 합성곱층과 풀링층을 통해 얻은 특징들을 기반으로 최종 출력값을 도출한다.

[0084] 장단기 기억 신경망(LSTM)은 데이터 길이가 길어짐에 따라 과거 정보가 손실되는 장기 의존성 문제를 해결하기 위해 제안된 순환 신경망이다. 장단기 기억 신경망은 기존 순환 신경망에 메모리 셀을 추가함으로써 장기 의존성 문제를 극복하였다. 메모리 셀은 망각 게이트, 입력 게이트, 출력 게이트로 구성되어 있으며 3개의 게이트를 통해 정보의 흐름을 조절하고 장기적인 정보를 유지한다. 메모리 셀의 구조는 도 4와 같다.

[0085] 첫 번째로 망각 게이트는 장기 기억 중 어떤 과거 정보를 기억할 것인지, 삭제할 것인지 결정하는 게이트이다. 수학식 5와 같이 이전 시점 은닉층 정보와 현재 입력값을 기반으로 시그모이드 함수를 활용하여 0과 1 사이의

값을 도출하고 과거 정보를 얼마나 반영할지 결정한다.

수학식 5

[0086] $$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

[0087] f_t는 망각 게이트를 의미하고, h_{t-1}, x_t는 각각 이전 은닉층의 상태 정보와 현재 입력값을 의미한다. σ는 시그모이드 함수, W_x와 b_x는 각 게이트의 가중치와 편향을 나타낸다.

[0088] 입력 게이트는 현재 입력값과 이전 셀 상태를 고려하여 어떤 정보를 새롭게 받아들일 것인지 결정하는 게이트이다. 수학식 6과 수학식 7과 같이 시그모이드 함수를 기반으로 현재 정보를 얼마나 반영할 것인지 결정하고 하이퍼볼릭 탄젠트 함수를 통해 셀 상태에 반영할 값을 결정한다. 최종셀 상태는 수학식 8과 같이 갱신된다.

수학식 6

[0089] $$i_t = \sigma(W_i[h_{t-1}, x_t] + b_f)$$

수학식 7

[0090] $$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C)$$

수학식 8

[0091] $$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

[0092] i_t, $\tilde{C}_t$는 입력 게이트와 현재 셀 상태를 나타내고, tanh는 하이퍼볼릭 탄젠트 함수를 의미한다.

[0093] 마지막 출력 게이트는 앞의 두 게이트와 같은 방식으로 출력값을 결정하고 이는 수학식 9와 같다. 수학식 10과 같이 메모리 셀의 최종 결괏값은 출력 게이트의 출력값과 갱신된 셀 상태의 하이퍼볼릭 탄젠트 함수 결괏값의 곱셈 연산을 통해 결정되며, 이러한 곱셈 연산을 통해 정보의 양을 조절할 수 있다.

수학식 9

[0094] $$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

수학식 10

[0095] $$h_t = o_t \circ \tanh(c_t)$$

[0096] 양방향 장단기 기억 신경망(BILSTM)은 순방향 장단기 기억 신경망에 역방향 장단기 기억 신경망을 추가하여 예측 성능을 향상시킨 모델이다. 시계열 데이터를 양방향으로 학습함으로써 과거 정보와 미래 정보에 대한 양방향 장기적 종속성을 학습한다. 양방향 장단기 기억 신경망은 에너지 예측을 위해 활용되고 있으며 구조는 도 5와

같다.

[0097] 서로 다른 인공신경망을 결합하여 성능을 향상시키는 혼합신경망 구조는 최근 이미지 분류, 행동 인지, 시계열 예측 등 많은 분야에서 좋은 성능을 보인다. 본 발명에서는 최근 에너지 예측 분야에서 많이 사용되는 CNN-LSTM과 CNN-BILSTM을 구현하였다. CNN은 합성 곱 연산을 통해 파라미터 수를 감소시킴과 동시에 공간적 특성을 추출하고 LSTM과 BILSTM은 이를 학습함으로써 공간적, 시간적 특성을 동시에 반영한 예측 결과를 도출한다.

[0099] 실험 데이터 및 실험 설정

[0100] 본 발명은 예측 모델의 학습과 평가를 위해 2019년 12월부터 2021년 7월까지 1시간 간격으로 계측된 실제 인천대학교 내 세 건물의 전력 사용량 데이터와 한국 기상청에서 제공하는 기상 데이터를 수집하였다. 3,920개의 데이터가 모델 학습을 위해 사용되었으며, 계절에 따른 예측 성능을 비교하기 위해 4, 8, 12, 2월에 해당하는 데이터 560개의 데이터가 테스트 과정에 사용되었다.

표 2

Table 2. Example of data on day n, time l

Variable	Value
Temperature	27.5℃
Humidity	100%
Wind Speed	$50 m/s$
Insolation	$0.54k\,Wh/m^2$
Cloudiness Index	10
Season Index	1
Weekday / Weekend	1
Semester / Vacation	1
Electricity Demand	$213 Kw$
Peak Time	11:00-15:00

[0101]

[0102] 표 2는 예측 모델의 입력 변수로 사용되는 데이터 예시를 나타낸다. 기후 정보와 전력 사용량은 측정된 수치를 사용하였으며 습도와 전운량은 각각 0과 100, 0과 10 사이의 범위에서 높을수록 습도와 운량이 높은 날씨임을 나타낸다. 계절은 계절별 상수를 선언하여 계절 정보를 나타내었으며 평일과 주말, 학기와 방학은 이진 변수로 구분하여 건물의 운영 정보를 나타내었다. 예측 대상으로 고려되는 피크 시간은 지역, 계절, 전력 사용 패턴 등에 따라 정의가 변화한다. 따라서, 기존에 여러 정의가 존재하나 본 발명에서는 수집한 데이터를 기반으로 전력 수요량이 가장 집중되는 11:00시부터 15:00까지를 피크시간으로 정의하였다.

[0103] 실험은 도 6과 같이 전력 사용량 패턴의 차이가 존재하는 세 건물을 대상으로 수행되었다. 도 6은 각 건물의 한 달간의 전력 수요 패턴을 나타낸다. 주로 강의실로 구성된 건물 C는 강의 스케줄에 따라 유동 인구의 변동성이 큰 만큼 다른 건물에 비해 가장 높은 변동성을 보인다. 이에 반해 연구실로 구성된 건물 B는 전반적으로 가장 낮은 전력 사용량을 보이며 변동성 역시 작다. 행정 관리실로 구성된 건물 A의 경우, 평균 전력 사용량은 가장 높지만, 건물 C에 비해 낮은 변동성을 보인다.

[0105] 예측 성능 평가 지표

[0106] 예측 성능을 평가하기 위해 본 발명은 기존 전력 예측 연구에서 널리 사용되는 평가지표인 절대 평균오차(MAE), 평균 제곱근 오차(RMSE), 평균 절대오차(MAPE)를 측정하였다. 각 평가 지표는 값이 작을수록 예측 성능이 우수함을 의미하며 수학식 11 내지 수학식 13과 같이 계산된다.

수학식 11

[0107] $$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\hat{y_t} - y_t)^2}$$

수학식 12

[0108] $$MAPE(\%) = \left(\frac{1}{N}\sum_{t=1}^{N}\left(\frac{|\hat{y_t} - y_t|}{y_t}\right)\right) \cdot 100$$

수학식 13

[0109] $$MAE = \frac{1}{N}\sum_{t=1}^{N}|y - \hat{y}|$$

[0110] 수학식 11 내지 수학식 13에서 N은 예측한 데이터의 개수이며, $\hat{y_t}$는 t 시점의 예측된 전력 사용량, y_t는 t 시점의 실제 전력 사용량을 의미한다.

[0112] 하이퍼 파라미터

[0113] 기계 학습과 딥러닝 모델은 일반적으로 하이퍼 파라미터에 따라 성능의 차이가 존재한다. 그러나 각 모델의 모든 하이퍼 파라미터 조합을 고려할 경우, 경우의 수가 매우 많으며 많은 시간이 소요된다. 따라서 본 발명에서는 충분한 반복 실험을 통해 모델별 가장 우수한 예측 성능을 보이는 하이퍼 파라미터를 설정하였다.

[0115] 실험 결과

[0116] 건물에 따른 전력 수요 예측 모델들의 MAE, MAPE, RMSE 결과는 표 3과 같다. 표 3은 각 건물에 대한 예측 모델들의 예측 성능을 나타내며 굵은 글씨는 가장 우수한 성능을 의미한다. 높은 전력 사용량과 변동성을 보이는 건물 A에서는 MAE, MAPE 관점에서 BiLSTM이 가장 우수한 예측 성능을 보였으며, RMSE 관점에서는 LSTM이 가장 낮은 예측 오차를 보였다. 이에 반해 전력 수요 규모가 가장 작은 건물 B에서는 CNN이 MAE와 RMSE 관점에서 우수한 성능을 보였으며, GB가 MAPE 관점에서 가장 우수한 성능을 보였다. 전력 수요의 변동성이 가장 높은 건물 C의 경우, MAE, MAPE, RMSE 관점에서 모두 BiLSTM이 가장 우수한 예측 성능을 보였다.

표 3

Table. 3. Comparison of electricity demand prediction model's performance in terms of *MAE, MAPE, RMSE*. (Bold indicate the model with best performance)

Building No.	Evaluation Index	GB	CAT	XGB	DNN	CNN	LSTM	BiLSTM	CNN-LSTM	CNN-BiLSTM
1	MAE	15.629	17.62	15.229	15.639	16.353	14.495	**14.351**	26.28	23.359
	MAPE	6.868	7.714	6.615	6.903	7.163	6.463	**6.379**	11.157	10.198
	RMSE	22.224	24.113	21.449	21.961	22.51	**20.613**	20.845	34.881	30.273
5	MAE	10.9	11.205	12.428	9.793	**9.054**	12.28	9.795	21.001	19.974
	MAPE	**12.719**	12.872	14.994	15.649	14.134	14.567	12.932	31.308	30.779
	RMSE	16.264	16.738	17.63	13.959	**13.157**	18.124	14.517	27.507	25.774
12	MAE	15.914	16.159	15.81	16.279	41.625	16.299	**15.485**	24.642	19.257
	MAPE	7.66	7.515	7.588	7.775	18.632	7.675	**7.434**	13.272	9.735
	RMSE	22.256	23.332	22.357	22.716	77.379	22.55	**22.098**	34.648	27.055

[0117]

[0118] 예측 결과에 따르면, 단일 모델이 모든 건물에 있어 가장 우수한 예측 성능을 보이지 않고 건물별 예측 성능의 차이가 존재함을 확인할 수 있다.

[0119] 이러한 성능 차이는 건물별 적절한 예측 모델의 선정을 통해 예측 정확도를 향상시킬 수 있음을 의미한다.

[0120] 따라서, 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법에 의하면, BILSTM을 사용하여, 평균 전력 사용량이 가장 높지만 강의실에 비해 낮은 변동성을 보이는 행정 관리실의 전력 수요를 예측하고, CNN을 사용하여, 낮은 전력 사용량을 보이며 변동성이 작은 연구실의 전력 수요를 예측하며, BILSTM을 사용하여, 강의 스케줄에 따라 유동 인구의 변동성이 커 가장 높은 변동성을 보이는 강의실의 전력 수요를 예측하는 것이 바람직할 수 있다.

[0121] 피크 시간 이전의 전력 수요와 피크 시간의 전력 수요 변동성에 따른 예측 성능을 분석하기 위해 본 발명은 4가지 사례에 대해 모델별 예측 성능을 비교하였으며 결과는 도 8과 같다. (a)와 (b)는 각각 피크 시간 이전과 피크 시간의 전력 수요 평균의 차이가 작고 피크 시간 내 전력 수요 변동성이 낮고 높은 사례를 나타낸다. (c)와 (d)는 각각 피크 시간 이전과 피크 시간의 전력 수요 평균의 차이가 크고 변동성이 낮고 높은 사례를 나타낸다.

[0122] (a)의 경우, BiLSTM이 다른 예측 모델에 비해 피크 시간의 전력 사용량 규모와 패턴을 잘 예측하였고 (b)의 경우, CATBoost가 가장 우수한 예측 성능을 보였다. (c)의 경우, (a)와 동일하게 BiLSTM이 가장 우수한 예측 성능을 보임에 따라 피크 시간의 변동성이 낮을 때, BiLSTM의 예측 성능이 우수함을 확인할 수 있다. 마지막으로 (d) 사례에서는 DNN이 가장 우수한 성능을 보였다.

[0123] 계절에 따른 예측 모델의 성능 차이를 분석하기 위해 본 발명은 도 9와 같이 각 건물의 계절별 MAE 분포를 박스플롯으로 나타내었다. 모든 건물에 있어 전력 수요가 높은 여름에 예측 오차가 크고 분산이 높은 경향을 보였고 봄과 가을에 예측 오차의 평균이 낮은 경향을 보였다. 또한, 건물 A와 C에서 LSTM과 BiLSTM과 같은 순환 신경망 기반의 예측 모델이 우수한 예측 성능을 보이는 반면 건물 B에서는 CNN 기반의 예측 모델이 가장 낮은 예측 오차를 보임을 확인할 수 있다.

[0124] 마지막으로 학습 데이터 양에 따른 예측 모델의 성능을 분석하기 위해 학습 데이터 4, 8, 12개월로 증가시키며 실험을 진행하였고 그 결과는 표 4와 같다. 건물 A와 B는 학습 데이터양이 8개월일 때, 대부분의 예측 모델이 가장 우수한 성능을 보였으며 건물 C의 경우, 학습 데이터양이 12개월일 때 가장 낮은 예측 오차를 보였다. 따라서 학습 데이터의 양은 예측 성능과 비례하지 않으며 적절한 학습 데이터양 설정을 통해 예측 성능을 향상시킬 수 있음을 확인할 수 있다.

표 4

Table. 4. Comparison of prediction model results based on amount of training data.

Building	Model	Amount for Training Data								
		4			8			12		
		MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE
Building A	GB	15.62	7.86	21.62	**14.09**	6.99	**20.05**	15.63	**6.87**	22.22
	XGBoost	16.39	8.39	22.80	**15.74**	**7.72**	**22.04**	17.62	7.71	24.11
	CATBoost	14.92	7.52	20.64	**12.95**	**6.52**	**18.64**	15.23	6.62	21.45
	DNN	14.56	7.13	20.03	**13.80**	**6.91**	**19.54**	15.64	6.90	21.96
	1D CNN	**14.67**	7.28	**19.82**	14.78	7.32	19.87	16.35	**7.16**	22.51
	LSTM	13.31	6.61	19.06	**12.99**	6.53	**18.45**	14.50	**6.46**	20.61
	BILSTM	13.01	6.46	18.75	**12.62**	**6.34**	**18.20**	14.35	6.38	20.85
	CNN + LSTM	**19.88**	**9.71**	**27.04**	22.03	11.07	29.47	26.28	11.16	34.88
	CNN + BILSTM	16.81	**8.15**	23.30	**16.72**	8.23	**22.86**	23.36	10.20	30.27
Building B	GB	9.51	18.84	12.95	**6.66**	13.78	**9.63**	10.90	**12.72**	16.26
	XGBoost	9.31	16.26	13.46	**6.69**	**12.27**	**10.14**	11.21	12.87	16.74
	CATBoost	19.69	43.14	27.86	**7.17**	**14.57**	**10.00**	12.43	14.99	17.63
	DNN	18.93	36.12	25.63	**8.98**	**15.73**	**12.04**	9.79	15.65	13.96
	1D CNN	37.55	43.08	66.87	**7.99**	**13.23**	**11.71**	9.05	14.13	13.16
	LSTM	30.77	59.32	43.97	**7.15**	**13.18**	**10.26**	12.28	14.57	18.12
	BILSTM	28.50	63.79	35.70	**6.97**	13.70	**10.14**	9.80	**12.93**	14.52
	CNN + LSTM	20.87	45.72	29.92	**13.41**	**28.12**	**19.38**	21.00	31.31	27.51
	CNN + BILSTM	26.94	66.66	36.79	**15.62**	**29.99**	**22.93**	19.97	30.78	25.77
Building C	GB	19.43	9.81	25.37	17.07	8.31	23.21	**15.91**	**7.66**	**22.26**
	XGBoost	20.62	9.79	28.11	17.37	8.17	23.90	**16.16**	**7.52**	**23.33**
	CATBoost	19.55	10.19	25.64	16.55	8.03	23.06	**15.81**	**7.59**	**22.36**
	DNN	18.56	9.11	25.06	17.49	8.74	24.21	**16.28**	**7.78**	**22.72**
	1D CNN	46.83	20.42	83.72	71.20	30.66	111.40	**41.63**	**18.63**	**77.38**
	LSTM	20.65	10.28	27.65	16.89	8.20	22.97	**16.30**	**7.68**	**22.55**
	BILSTM	19.86	9.92	26.35	15.75	7.51	**22.03**	**15.49**	**7.43**	22.10
	CNN + LSTM	33.90	18.83	48.07	**24.12**	**12.93**	**32.06**	24.64	13.27	34.65
	CNN + BILSTM	28.43	16.42	39.61	24.76	13.07	34.26	**19.26**	**9.74**	**27.06**

[0125]

[0126] 본 명세서에서 설명된 위 실시예 및 도면들은 단지 예시적인 것일 뿐, 어떠한 방법으로도 본 발명의 범위를 한정하는 것은 아니다. 또한, 도면에 도시된 구성요소들 간의 선들의 연결 또는 연결 부재들은 기능적인 연결 및/또는 물리적 또는 회로적 연결들을 예시적으로 나타낸 것으로서, 실제장치에서는 대체 가능하거나 추가의 다양한 기능적인 연결, 물리적인 연결, 또는 회로 연결들로서 나타내어질 수 있다. 또한, "필수적인", "중요하게" 등과 같이 구체적인 언급이 없다면 본 발명의 적용을 위하여 반드시 필요한 구성요소가 아닐 수 있다.

[0127] 상기의 상세한 설명은 모든 면에서 제한적으로 해석되어서는 아니되고 예시적인 것으로 고려되어야 한다. 본 발명의 범위는 첨부된 청구항의 합리적 해석에 의해 결정되어야 하고, 본 발명의 등가적 범위 내에서의 모든 변경은 본 발명의 범위에 포함된다.

[0128] 본 발명의 명세서(특히 특허청구범위에서)에서 "상기"의 용어 및 이와 유사한 지시 용어의 사용은 단수 및 복수 모두에 해당하는 것일 수 있다. 또한, 본 발명에서 범위(range)를 기재한 경우 상기 범위에 속하는 개별적인 값을 적용한 발명을 포함하는 것으로서(이에 반하는 기재가 없다면), 발명의 상세한 설명에 상기 범위를 구성하는 각 개별적인 값을 기재한 것과 같다. 또한, 본 발명 중 방법 발명에서 제시하는 단계들은 반드시 그 선후의 순서에 대한 구속을 의도한 것이 아니며, 각 공정의 본질에 따라 반드시 어느 단계가 선행되어야 하는 것이 아닌 한 순서는 필요에 따라 적절히 변경될 수 있다. 본 발명에서 모든 예들 또는 예시적인 용어(예들 들어, 등등)의 사용은 단순히 본 발명을 상세히 설명하기 위한 것으로서 특허청구범위에 의해 한정되지 않는 이상 상기 예들 또는 예시적인 용어로 인해본 발명의 범위가 한정되는 것은 아니다. 또한, 통상의 기술자는 다양한 수정, 조합 및 변경이 부가된 특허청구범위 또는 그 균등 범주 내에서 설계 조건 및 요소에 따라 구성될 수 있음을 이해할 수 있다.

부호의 설명

[0129] 110 : 훈련 데이터 세트 112 : 테스트 데이터 세트

900 : 본 발명의 일 실시예에 의한 건물별 맞춤형 전력 수요 예측 모델 선정을 위한 예측 모델 성능 비교 방법

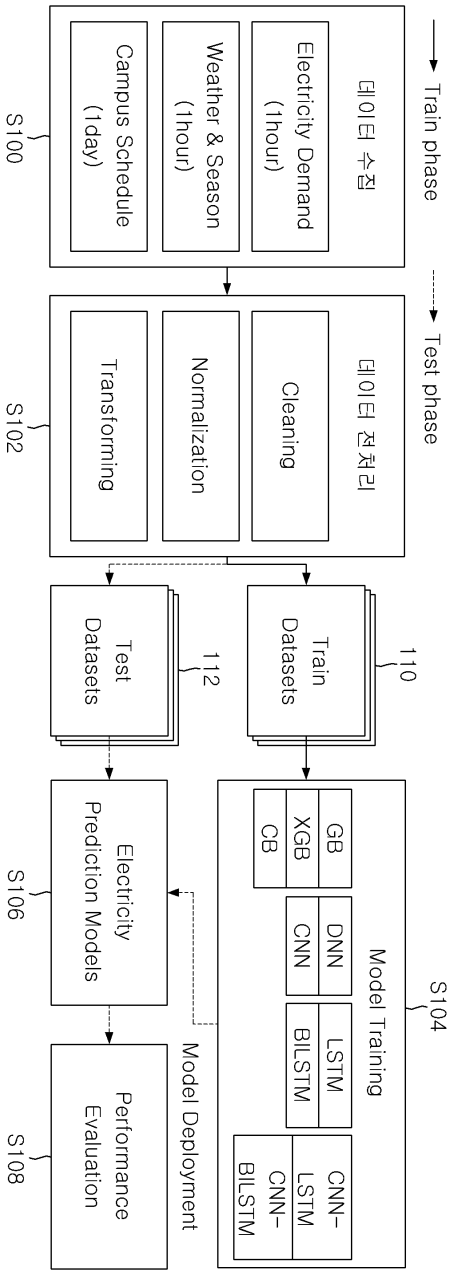
이 수행될 수 있는 장치

902 : 프로세서 904 : 메모리

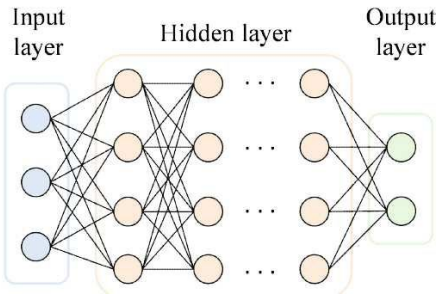
906 : 네트워크 통신 인터페이스 908 : 입출력 인터페이스

도면

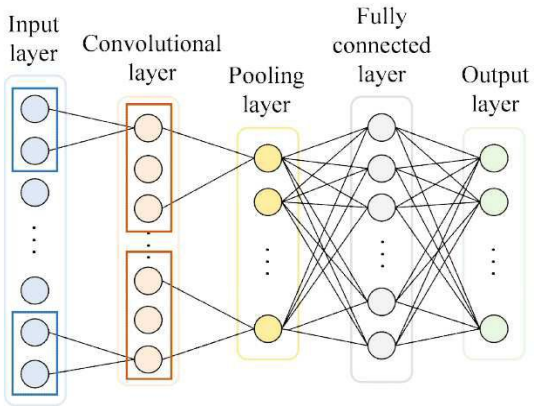
도면1



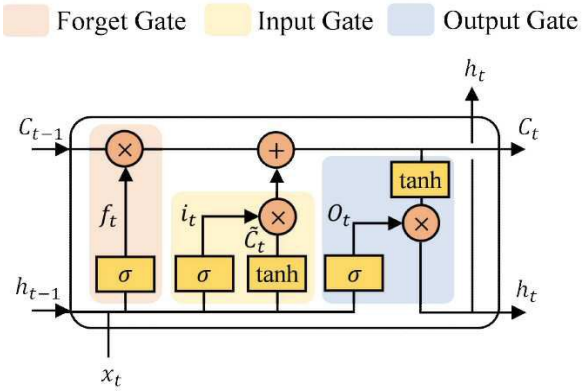
도면2



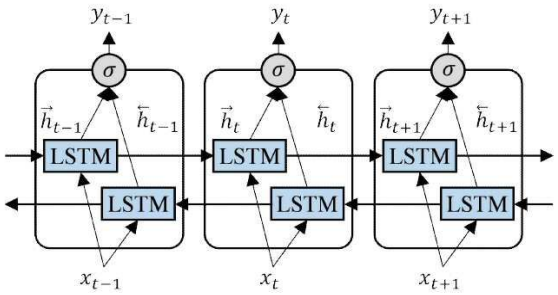
도면3



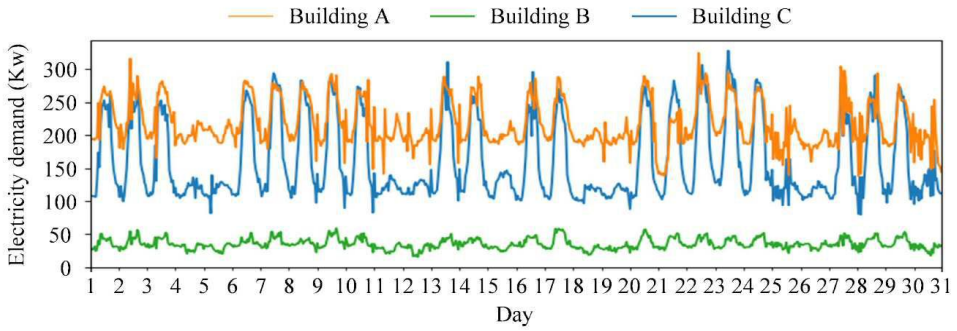
도면4



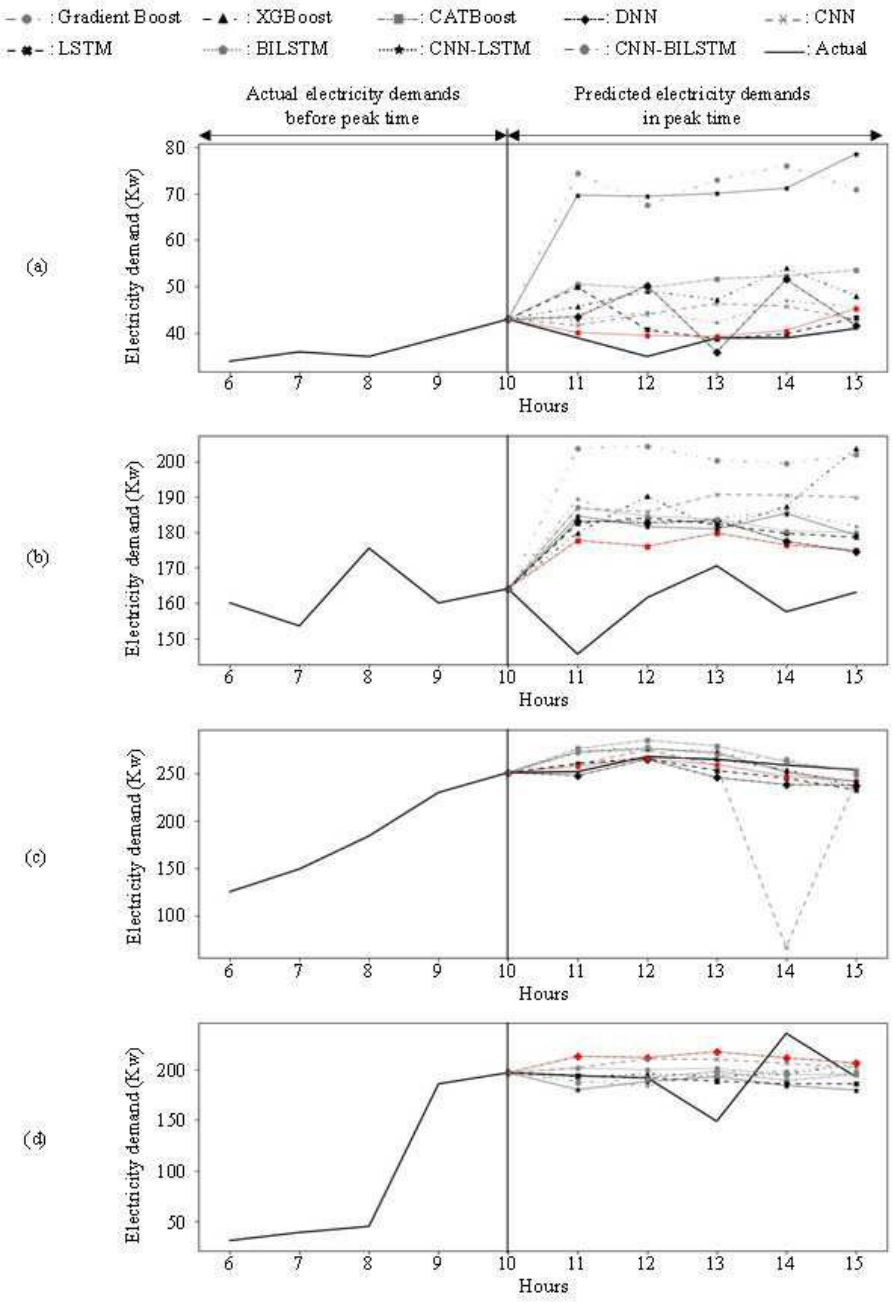
도면5



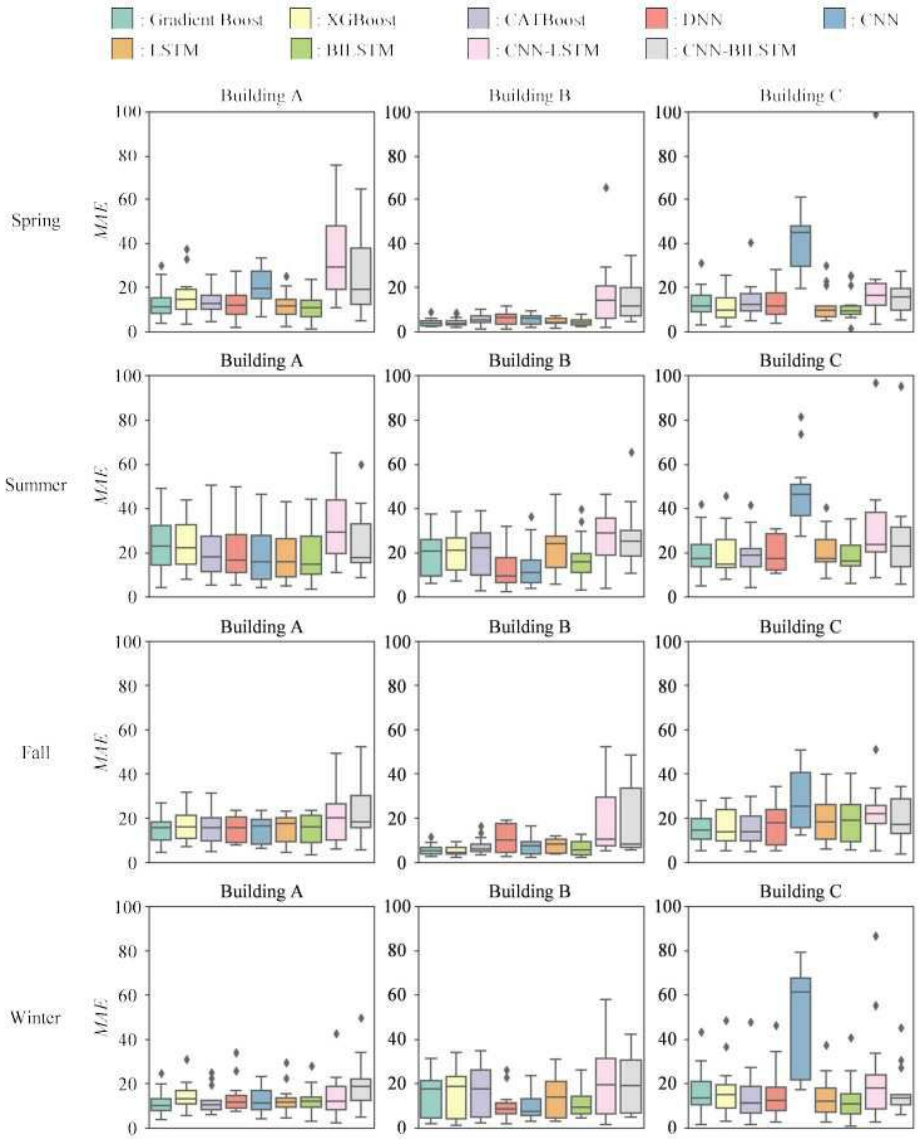
도면6


Building A
Building B
Building C
Electricity demand (Kw)
300
250
200
150
100
50
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Day

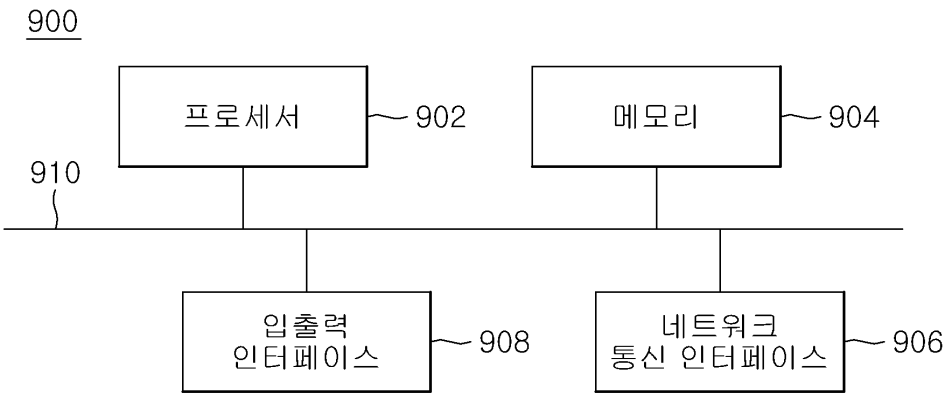
도면7


: Gradient Boost
: XGBoost
: CATBoost
: DNN
: CNN
: LSTM
: BILSTM
: CNN-LSTM
: CNN-BILSTM
: Actual
Actual electricity demands before peak time
Predicted electricity demands in peak time
(a)
(b)
(c)
(d)
Electricity demand (Kw)
Hours

도면8


: Gradient Boost
: XGBoost
: CATBoost
: DNN
: CNN
: LSTM
: BILSTM
: CNN-LSTM
: CNN-BILSTM
Building A
Building B
Building C
Spring
Summer
Fall
Winter
MAE
0
20
40
60
80
100

도면9


900
프로세서
902
메모리
904
910
입출력
인터페이스
908
네트워크
통신 인터페이스
906