(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0002889
(43) 공개일자 2024년01월08일

(51) 국제특허분류(Int. Cl.)
H02S 50/00 (2014.01) *G06F 18/214* (2023.01)
G06F 30/27 (2020.01) *G06N 3/08* (2023.01)
(52) CPC특허분류
H02S 50/00 (2013.01)
G06F 18/214 (2023.01)
(21) 출원번호 10-2022-0182415
(22) 출원일자 2022년12월23일
심사청구일자 2022년12월23일
(30) 우선권주장
1020220080801 2022년06월30일 대한민국(KR)

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
황인준
서울특별시 성동구 매봉길 15(옥수동, 래미안 옥수 리버젠)
김동준
서울특별시 서초구 잠원로14길 23(잠원동, 롯데캐슬갤럭시)
(뒷면에 계속)
(74) 대리인
특허법인주연케이알피

전체 청구항 수 : 총 10 항

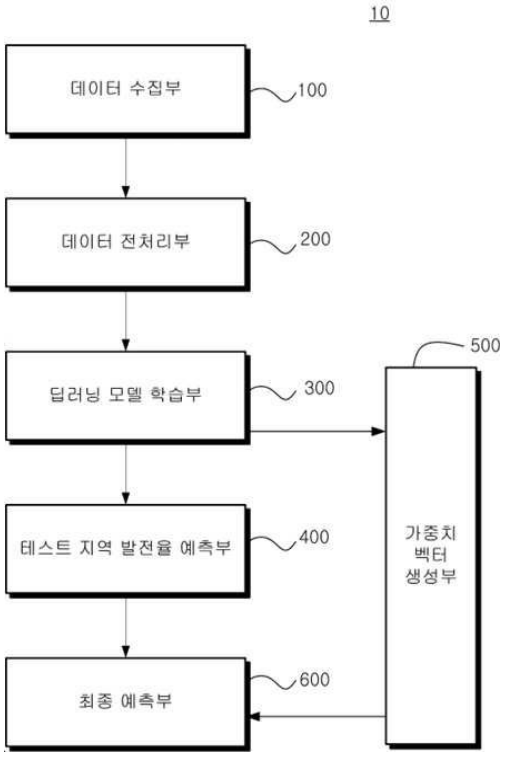
(54) 발명의 명칭 **딥러닝 모델 기반 태양광 발전율 예측 장치 및 이의 동작 방법**

(57) 요 약

본 출원의 실시 예에 따른 태양광 발전율 예측 장치는, 복수의 훈련 지역에 대해, 태양광 발전량 데이터, 기상관측 데이터 및 대기질 특정 데이터를 시계열적으로 수집하는 데이터 수집부; 상기 데이터 수집부가 수집한 데이터에 대해, 표준 정규화를 수행하는 데이터 전처리부; 상기 훈련 지역 각각의 데이터를 이용하여, 테스트 지역의

(뒷면에 계속)

대 표 도 - 도1



발전율 예측에 필요한 딥러닝 모델을 학습하는 딥러닝 모델 학습부; 학습된 상기 딥러닝 모델에 상기 테스트 지역의 입력 변수를 활용하여 발전율 예측값 벡터를 생성하는 테스트 지역 발전율 예측부; 상기 훈련 지역과 상기 테스트 지역의 기상 및 대기질 유사도를 기반으로 가중치 벡터를 생성하는 가중치 벡터 생성부; 및 상기 발전율 예측값 벡터와 상기 가중치 벡터를 활용하여 상기 테스트 지역의 발전율을 예측하는 최종 예측부를 포함한다. 본 출원의 실시 예에 따른 태양광 발전율 예측 장치는 기상 유사도에 기반한 가중치를 산출하여 예측을 수행함으로써 높은 정확도와 해석이 용이한 예측값을 제공할 수 있다.

(52) CPC특허분류
G06F 30/27 (2020.01)
G06N 3/049 (2023.01)
G06N 3/08 (2023.01)
Y02E 10/50 (2020.08)

(72) 발명자
박성우
서울특별시 성동구 상원12길 15(성수동1가, 성수동 동아그린아파트)

문재욱
서울특별시 광진구 아차산로78길 7(광장동, 유진스웰)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711157798
과제번호	2019M3F2A1073184
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	에너지클라우드기술개발(과기정통부)
연구과제명	개방형 마이크로그리드 플랫폼을 위한 딥러닝 기반의 확률론적 에너지 사용량 예측 모델 구성 및 시뮬레이션 개발
기 여 율	1/1
과제수행기관명	고려대학교
연구기간	2022.01.01 ~ 2022.12.31

명 세 서

청구범위

청구항 1

복수의 훈련 지역에 대해, 태양광 발전량 데이터, 기상관측 데이터 및 대기질 특정 데이터를 시계열적으로 수집하는 데이터 수집부;

상기 데이터 수집부가 수집한 데이터에 대해, 표준 정규화를 수행하는 데이터 전처리부;

상기 훈련 지역 각각의 데이터를 이용하여, 테스트 지역의 발전율 예측에 필요한 딥러닝 모델을 학습하는 딥러닝 모델 학습부;

학습된 상기 딥러닝 모델에 상기 테스트 지역의 입력 변수를 활용하여 발전율 예측값 벡터를 생성하는 테스트 지역 발전율 예측부;

상기 훈련 지역과 상기 테스트 지역의 기상 및 대기질 유사도를 기반으로 가중치 벡터를 생성하는 가중치 벡터 생성부; 및

상기 발전율 예측값 벡터와 상기 가중치 벡터를 활용하여 상기 테스트 지역의 발전율을 예측하는 최종 예측부를 포함하는 태양광 발전율 예측 장치.

청구항 2

제1 항에 있어서,

상기 데이터 전처리부는,

상기 태양광 발전량 데이터를 각 발전소의 설비용량으로 나눈 발전율로 변환하여 상기 발전율을 출력변수로 활용하는 태양광 발전율 예측 장치.

청구항 3

제1 항에 있어서,

상기 딥러닝 모델 학습부는,

상기 훈련 지역에서 각각 상기 딥러닝 모델의 학습을 수행하고, 상기 훈련 지역의 개수만큼의 상기 딥러닝 모델을 구축하는 태양광 발전율 예측 장치.

청구항 4

제1 항에 있어서,

상기 딥러닝 모델 학습부는,

상기 딥러닝 모델로 2-layer DNN(Deep Neural Network)을 활용하는 태양광 발전율 예측 장치.

청구항 5

제1 항에 있어서,

상기 가중치 벡터 생성부는,

상기 훈련 지역 중 i번째 지역에 대해서 발전율과 상기 기상관측 데이터 및 상기 대기질 측정 데이터와 관련된 V개 변수 간의 상관관계를 도출한 크기 $1\times V$의 벡터를 생성하고,

상기 i번째 지역과 상기 테스트 지역 각각의 상기 기상관측 데이터 및 상기 대기질 측정 데이터와 관련된 변수의 상관관계를 도출한 크기 $1\times V$의 벡터를 생성하는 태양광 발전율 예측 장치.

청구항 6

제5 항에 있어서,

상기 가중치 벡터 생성부는,

상기 크기 $1\times V$의 벡터 두개를 내적곱하고, 내적곱한 결과 벡터의 요소의 합이 1이 되도록 조정하는 태양광 발전율 예측 장치.

청구항 7

태양광 발전율 예측 장치에 의한 동작 방법으로서,

복수의 훈련 지역에 대해, 태양광 발전량 데이터, 기상관측 데이터 및 대기질 특정 데이터를 시계열적으로 수집하는 단계;

수집한 데이터에 대해, 표준 정규화를 수행하는 단계;

테스트 지역의 발전율 예측에 필요한 딥러닝 모델을 학습하는 단계;

학습된 상기 딥러닝 모델에 상기 테스트 지역의 입력 변수를 활용하여 발전율 예측값 벡터를 생성하는 단계;

상기 훈련 지역과 상기 테스트 지역의 기상 및 대기질 유사도를 기반으로 가중치 벡터를 생성하는 단계; 및

상기 발전율 예측값 벡터와 상기 가중치 벡터를 활용하여 상기 테스트 지역의 발전율을 예측하는 단계를 포함하는 태양광 발전율 예측 방법.

청구항 8

제7 항에 있어서,

상기 표준 정규화를 수행하는 단계에서,

상기 태양광 발전량 데이터를 각 발전소의 설비용량으로 나눈 발전율로 변환하여 상기 발전율을 출력변수로 활용하는 단계를 더 포함하는 태양광 발전율 예측 방법.

청구항 9

제7 항에 있어서,

상기 딥러닝 모델을 학습하는 단계에서,

상기 훈련 지역에서 각각 상기 딥러닝 모델의 학습을 수행하고, 상기 훈련 지역의 개수만큼의 상기 딥러닝 모델을 구축하는 태양광 발전율 예측 방법.

청구항 10

제7 항에 있어서,

상기 가중치 벡터를 설정하는 단계에서,

상기 훈련 지역 중 i번째 지역에 대해서 발전율과 상기 기상관측 데이터 및 상기 대기질 측정 데이터와 관련된 V개 변수 간의 상관관계를 도출한 크기 $1\times V$의 벡터를 생성하고,

상기 i번째 지역과 상기 테스트 지역 각각의 상기 기상관측 데이터 및 상기 대기질 측정 데이터와 관련된 변수의 상관관계를 도출한 크기 $1\times V$의 벡터를 생성하며,

상기 크기 $1\times V$의 벡터 두개를 내적곱하고, 내적곱한 결과 벡터의 요소의 합이 1이 되도록 조정하는 태양광 발전율 예측 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 딥러닝 모델 기반 태양광 발전율 예측 장치 및 이의 동작 방법에 관한 것이다.

배 경 기 술

[0002] 석탄, 석유 등을 원료로 하는 발전 방식은 저렴한 단가로 인해 국내외에서 가장 많이 활용되는 발전 방식이지만, 보건 및 환경 문제를 유발하는 다량의 유해 물질들은 배출한다는 단점 때문에 최근 이를 대체할 발전원으로 신재생 에너지가 대두되고 있다. 그 중에서도 태양광 발전은 설치 및 유지보수가 용이하다는 장점 때문에 가장 높은 비중과 성장률을 보이고 있다.

[0003] 태양광 발전소의 안정적인 운영 및 전력 공급을 위해서는 발전량의 예측을 통해 변동성이 강한 태양광 에너지의 불확실성을 개선하고 전력 공급의 신뢰성을 향상해야 한다. 이를 위하여 태양광 발전소에서 측정된 태양광 발전량 정보와 발전소가 위치한 지역의 기상 정보를 사용한 기계학습 및 심층학습 기반의 태양광 발전량 예측 연구가 활발하게 진행되고 있다. 이러한 연구들은 좋은 성능을 보여주고 있으나 예측 모델의 학습을 위해서는 많은 양의 과거 데이터가 필요하다. 만약 과거 데이터가 충분하지 않으면 모델의 예측 성능이 낮게 나올 수 있으며 이를 콜드스타트 문제라고 한다.

[0004] 신규 태양광 발전소가 증가하고 있음에도 불구하고, 태양광 발전량 예측의 콜드스타트 문제를 해결할 수 있는 방법에 관한 연구는 아직까지 미비한 실정이다. 새롭게 건설할 태양광 발전소에 대한 정확한 발전량 예측이 가능해진다면 건설 추진 단계에서 사업성 및 건설 타당성을 평가하기 용이해져서 보급 확대에 도움이 될 수 있다.

발명의 내용

해결하려는 과제

[0005] 본 발명의 목적은 예측 대상 지역에 과거 태양광 발전량 데이터가 없을 때 발전율 예측 모델 성능의 저하를 해결하기 위해 기상 유사도를 통해 유사한 지역의 과거 데이터를 활용하는 제로 샷 학습 기반의 태양광 발전율 예측 장치 및 이의 동작 방법을 제공하는데 있다.

과제의 해결 수단

[0006] 본 출원의 일 실시예에 따른 태양광 발전율 예측 장치는, 복수의 훈련 지역에 대해, 태양광 발전량 데이터, 기상관측 데이터 및 대기질 특정 데이터를 시계열적으로 수집하는 데이터 수집부; 상기 데이터 수집부가 수집한 데이터에 대해, 표준 정규화를 수행하는 데이터 전처리부; 상기 훈련 지역 각각의 데이터를 이용하여, 테스트 지역의 발전율 예측에 필요한 딥러닝 모델을 학습하는 딥러닝 모델 학습부; 학습된 상기 딥러닝 모델에 상기 테스트 지역의 입력 변수를 활용하여 발전율 예측값 벡터를 생성하는 테스트 지역 발전율 예측부; 상기 훈련 지역과 상기 테스트 지역의 기상 및 대기질 유사도를 기반으로 가중치 벡터를 생성하는 가중치 벡터 생성부; 및 상기 발전율 예측값 벡터와 상기 가중치 벡터를 활용하여 상기 테스트 지역의 발전율을 예측하는 최종 예측부를 포함한다.

[0007] 본 출원의 일 실시예에 따른 태양광 발전율 예측 방법은, 태양광 발전율 예측 장치에 의한 동작 방법으로서, 복수의 훈련 지역에 대해, 태양광 발전량 데이터, 기상관측 데이터 및 대기질 특정 데이터를 시계열적으로 수집하는 단계; 수집한 데이터에 대해, 표준 정규화를 수행하는 단계; 테스트 지역의 발전율 예측에 필요한 딥러닝 모델을 학습하는 단계; 학습된 상기 딥러닝 모델에 상기 테스트 지역의 입력 변수를 활용하여 발전율 예측값 벡터를 생성하는 단계; 상기 훈련 지역과 상기 테스트 지역의 기상 및 대기질 유사도를 기반으로 가중치 벡터를 생성하는 단계; 및 상기 발전율 예측값 벡터와 상기 가중치 벡터를 활용하여 상기 테스트 지역의 발전율을 예측하는 단계를 포함한다.

발명의 효과

[0008] 본 출원의 실시예에 따른 태양광 발전율 예측 장치 및 이의 동작 방법은 기상도가 유사한 지역의 과거 데이터를 활용하는 제로 샷 학습을 사용하여 발전율 정보가 없는 새로운 지역의 발전율 예측 모델 성능의 저하 문제를 개선할 수 있다.

도면의 간단한 설명

[0009] 도 1은 본 출원의 실시 예에 따른 태양광 발전율 예측 장치(10)를 보여주는 블록도이다.

도 2는 본 출원의 일 실시예에 따른 태양 효율 예측 장치의 구조를 나타낸 도면이다.

도 3a 및 도 3b에서 설명하는 내용은 본 출원의 일 실험예에 따른 평가 지표 중에서 가장 좋은 결과를 보인 진주 지역의 실험 결과 예시이다.

도 4는 본 출원의 일 실시예에 따른 태양광 발전율 예측 장치의 동작 방법을 보여주는 순서도이다.

도 5는 본 출원의 일 실시예에 따라 기상 유사도를 통하여 가중치를 구하는 과정을 나타내는 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0010] 이하에서는, 본 발명의 기술적 사상을 본 출원의 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있을 정도로 상세히 설명하기 위하여, 첨부된 도면을 참조하여 본 출원의 실시 예들이 자세히 설명될 것이다.

[0011] 도 1은 본 출원의 실시 예에 따른 태양광 발전율 예측 장치(10)를 보여주는 블록도이다.

[0012] 도 1을 참조하면, 태양광 발전량 예측 장치(10)는 과거에 측정된 발전량 정보가 풍부한 여러 지역으로부터 기상 정보와 태양광 발전량 정보 등을 수집하는 데이터 수집부(100), 발전량을 설비용량으로 나눈 발전율 값을 활용하여 표준 정규화를 통해 데이터 수집부(100)가 수집한 데이터를 전처리 하는 데이터 전처리부(200), 각 지역에 대한 태양광 발전율을 예측하는 딥러닝 모델을 각각 학습시키는 딥러닝 모델 학습부(300), 발전량 정보가 부족한 목표 지역의 기상 정보를 각 딥러닝 모델에 넣어 예측하는 테스트 지역 발전율 예측부(400), 예측된 값들에 각 딥러닝 모델이 학습된 지역의 기상 정보와 목표 지역의 기상 정보의 유사도를 통해 가중치를 주는 가중치 벡터 생성부(500) 및 가중치 값들로 부터 최종 발전율 예측값을 도출하는 최종 예측부(600)를 포함한다.

[0013] 데이터 수집부(100)는 태양광 발전율 예측 모델을 구성하기 위해 발전량 정보가 풍부한 여러 지역으로부터 기상 정보와 태양광 발전량 정보 등을 수집한다. 자연적인 가상현상 외에도 산업화로 인한 공장 매연, 부산물로 인한 (초)미세 먼지나 스모그 등도 태양광 변동성에 영향을 줄 수 있으므로 대기질에 관한 데이터도 함께 수집할 필요가 있다. 예를 들어 데이터 수집부(100)가 딥러닝 모델 학습용으로 수집한 데이터셋은 공공데이터 포털, 기상자료 개방 포털, 에어 코리아 세 곳에서 제공하는 데이터를 사용할 수 있다. 예를 들어, 공공데이터 포털에서는 전국에 위치한 태양광 발전소 중 진주, 제주, 부산, 영월, 인천, 하동, 신안 7곳의 태양광 발전량 데이터를 수집할 수 있다. 예를 들어, 기상자료 개방 포털 에서는 각 발전소와 가장 가까운 기상 관측소에서 측정된 기온, 습도, 풍속, 강수량, 증기압, 현지기압, 일조, 지면온도 데이터 등을 수집할 수 있다. 예를 들어, 에어 코리아에서는 각 발전소와 가장 가까운 대기질 측정소에서 측정된 미세먼지(PM10), 초미세먼지(PM25), 이산화질소(NO2), 일산화탄소(CO), 오존(O3), 아황산가스(SO2) 농도 데이터 등을 수집할 수 있다.

[0014] 데이터 수집부(100)는 데이터를 수집할 때에, 이에 대응하는 시간 데이터를 함께 수집할 수 있다. 태양광 발전량은 계절의 변화, 하루 중의 시간 변화의 영향을 많이 받는다. 예를 들어, 데이터 수집부(100)는 기상 데이터를 수집할 때에 월(month), 일(day), 시(hour) 및 이에 대한 시간 데이터를 1시간 단위로 함께 수집할 수 있다.

[0015] 데이터 전처리부(200)는 안정적인 딥러닝 모델 학습 및 범용성을 위해 전처리를 수행한다. 먼저 각 발전소에서 수집한 발전량 데이터는 각 발전소의 설비 용량으로 나누어 발전율 데이터로 변환하여 사용한다. 이는 처음 수집된 발전량을 변환없이 그대로 사용할 경우 각 발전소마다 설비 용량이 다르기 때문에 일관적인 딥러닝 모델 학습이 어렵기 때문이다. 발전율 데이터로 변환함으로써 기존에 존재하는 여러 지역의 발전소에서 학습되는 딥러닝 모델들을 통합적으로 활용 가능하다. 도출된 발전율 데이터와 데이터 수집부에서 수집한 기상관측 데이터, 대기질 측정 데이터에 대해서 표준 정규화(Standard Normalization)를 수행한다. 예를 들어, 기상 데이터 중에서 강수량과 일조 값의 경우, 비가 오지 않거나 구름이 많은 경우 측정값이 기록되지 않으므로, 두 기상 데이터의 결측값은 모두 0으로 대체하고, 다른 기상 데이터는 선형보간법으로 결측치를 채울 수 있다. 정규화를 통해 전처리가 끝난 데이터들은 각 훈련 지역별로 딥러닝 모델을 학습하는 입, 출력 데이터로 사용할 수 있다. 예를 들어, 정규화를 거쳐서 발전소가 위치하는 7개 지역 각각에 대해서 14개의 입력변수, 1개의 출력변수를 구성할 수 있다.

[0016] 딥러닝 모델 학습부(300)는 모델 앞선 데이터 전처리부(200)의 결과물로 출력된 발전율 데이터, 기상관측 데이터, 대기질 측정 데이터 등을 사용할 수 있다. 먼저 테스트 지역을 제외한 각 훈련 지역별로 동일한 딥러닝 모

델을 구축한다. 예를 들어, 데이터를 수집한 7개 지역 중 1개 지역을 발전율 데이터가 존재하지 않는다고 가정하는 테스트 지역으로 선정하고, 나머지 6개 지역은 발전율 데이터가 존재하는 훈련 지역으로 선정할 수 있다. 일 실시예에 따라, 딥러닝 모델로 2-layer DNN(Deep Neural Network)을 활용할 수 있다. 예를 들어, 각 훈련 지역별로 예측하고자 하는 시점으로부터 앞선 3일간의 기상관측 데이터, 대기질 측정 데이터는 입력 변수로 사용하고, 예측하고자 하는 시점의 발전율 데이터는 출력 변수로 사용하여 딥러닝 모델을 학습할 수 있다. 일 실시예에 따라 딥러닝 모델 학습부(300)는 훈련 지역 개수만큼의 딥러닝 모델을 구축할 수 있다.

[0017] 테스트 지역 발전율 예측부(400)는 이전 단계의 딥러닝 모델 학습부(300)에서 학습한 각 훈련 지역별 딥러닝 모델에 테스트 지역의 기상관측 데이터, 대기질 측정 데이터를 입력 변수로 하여 발전율을 예측한다. 일 실시예에 따라 테스트 지역 발전율 예측부(400)는 각 훈련 지역 개수 크기의 발전율 예측값 벡터를 생성할 수 있다.

[0018] 가중치 벡터 생성부(500)는 훈련 지역들과 테스트 지역의 기상 유사도를 반영한 가중치 벡터를 생성한다. 먼저 각 훈련 지역에서 입력 변수로 사용된 기상관측 데이터, 대기질 측정 데이터와 출력 변수로 사용된 발전율 데이터 간의 상관관계를 구한 행렬을 생성한다. 해당 행렬 생성함으로써 각 지역 별로 기상 변수들이 발전율에 끼치는 영향의 정도를 반영할 수 있다. 예를 들어, 상관관계를 구하는 방법으로는 PCC(Pearson Correlation Coefficient)를 활용할 수 있다. 예를 들어, 상관관계를 구하는 과정에서는 매일 오후 1시를 기준으로 측정된 데이터를 별개로 추출하여 시간에 따른 광량의 변동으로 인해 발생할 수 있는 오차를 줄일 수 있다.

[0019] 다음으로는 각 훈련 지역의 기상관측 데이터, 대기질 측정 데이터와 테스트 지역의 기상관측 데이터, 대기질 측정 데이터 간의 상관관계를 구한 행렬을 생성한다. 예를 들어, 상관관계는 동일하게 PCC로 구할 수 있다. 생성된 두 행렬을 지역별로 내적값을 도출한 벡터를 생성한 뒤, 벡터 값의 합이 1이 되도록 조정하여 최종 가중치 벡터를 생성한다. 기상 유사도를 통하여 가중치를 구하는 상세한 과정은 도 5에서 후술한다.

[0020] 최종 발전율 예측부(600)는 테스트 지역 발전율 예측부(400)에서 생성한 발전율 예측값 벡터와 가중치 벡터 생성부(500)에서 생성한 가중치 벡터를 내적곱하여 테스트 지역에 대한 최종적인 발전율 예측값을 도출한다. 예를 들어, 테스트 지역의 기상 정보를 학습된 모델들에 넣어서 도출한 예측값들을 하나로 합친 크기 $1\times N$의 벡터 Y_{new}'를 생성한다. 이후 기상 유사도를 통해 도출된 $1\times N$크기의 가중치 W_{new}'를 Y_{new}'와 곱하여 최종 발전율 예측값 Y_{new}를 구한다.

[0021] 상술한 바와 같이, 본 출원의 실시 예에 따른 태양광 발전율 예측 장치(10)는 지역별 기존 발전소들의 기상 변수와 발전율 값을 활용하여 심층학습 모델을 각각 학습하고, 테스트 지역의 기상 변수를 각 모델의 입력 변수로 하여 예측한 발전율 값들을 테스트 지역과 기존 발전소 지역과의 기상 유사도를 고려하여 최종적인 발전율 값을 예측한다. 따라서, 본 출원에 따른 태양광 발전율 예측 장치(10)는 새롭게 건설할 태양광 발전소에 대한 정확한 발전량 예측이 가능해지므로 건설 추진 단계에서 사업성 및 건설 타당성을 평가하기 용이해져서 태양광 발전소의 보급 확대에 도움이 될 수 있다.

[0022] 도 2는 본 출원의 일 실시예에 따른 태양 효율 예측 장치의 구조를 나타낸 도면이다.

[0023] 도 2를 참조하면, 먼저 데이터 수집부(100)에서 공공데이터 포털, 기상자료 개방 포털, 에어 코리아 등 세곳에서 공공데이터를 수집한다. 데이터 전처리부(200)에서는 발전소 별로 설비용량이 다른점을 보완하기 위해 발전량을 설비용량으로 나눈 발전율 값을 출력 변수로 활용한다. 또한 수집한 발전율 데이터, 기상관측 데이터, 대기질 측정 데이터의 범위가 각각 상이하므로 평균은 0으로, 표준 편차를 1로 정규화하는 표준 정규화를 수행한다.

[0024] 전처리가 끝난 데이터는 딥러닝 모델 학습부(300)로 전송된다. 데이터가 수집된 각 훈련 지역별로 동일한 딥러닝 모델을 구축하므로 예를 들어, N개의 훈련 지역에 따라 N개의 딥러닝 모델이 구축될 수 있다. 이 때, 기상관측 데이터, 대기질 측정 데이터는 입력변수로 활용하고, 발전율 데이터는 출력변수로 활용하여 딥러닝 모델을 학습한다. 일 실시예에 따라 딥러닝 모델로는 DNN을 활용할 수 있다.

[0025] 테스트 지역 발전율 예측부(400)에서는 딥러닝 모델 학습부(300)에서 학습한 훈련 지역 별 딥러닝 모델에 테스트 지역의 기상관측 데이터, 대기질 측정 데이터를 입력변수로 하여 테스트 지역의 발전율을 예측한다. 각 지역별 모델에서 도출된 발전율 예측값을 하나의 벡터로 이어붙이므로 훈련 지역 개수의 크기, 예를 들어 N개 훈련 지역에 따라 N개 성분에 따른 크기의 발전율 예측값 벡터가 생성된다.

[0026] 가중치 벡터 생성부(500)는 최종적인 발전율 예측값을 구하기 위해 발전율 예측값 벡터와 내적 연산하기 위한 가중치 벡터를 생성한다. 각 훈련 지역에서 입력변수로 사용된 기상관측 데이터, 대기질 측정 데이터와 출력 변

수로 사용된 발전율 데이터간의 상관관계를 구한 행렬을 생성한다. 상기 행렬에는 각 지역별로 기상 변수들이 발전율에 끼치는 영향의 정도를 반영할 수 있다. 더해서, 가중치 벡터 생성부(500)는 각 훈련 지역의 기상관측 데이터, 대기질 측정 데이터와 테스트 지역의 기상관측 데이터, 대기질 측정 데이터 간의 상관관계를 구한 행렬을 생성한다. 이 행렬에서는 각 훈련지역과 테스트 지역의 입력변수가 유사한 정도를 반영할 수 있다. 생성된 두 행렬을 지역별로 내적값을 도출하여 이어붙인 벡터를 생성하고 벡터 값의 합이 1이 되도록 조정하여 최종 가중치 벡터를 생성한다.

[0027] 최종 발전율 예측부(600)에서 테스트 지역 발전율 예측부(400)에서 생성한 발전율 예측값 벡터, 가중치 생성부(500)에서 생성한 가중치 벡터를 내적곱하여 테스트 지역에 대한 최종적인 발전율 예측값을 도출한다.

실 시 예 1

[0028] 도 3a 및 도 3b에서 설명하는 내용은 본 출원의 일 실험예에 따른 평가 지표 중에서 가장 좋은 결과를 보인 진주 지역의 실험 결과 예시이다. 도 3a는 다른 지역들과 기상 유사도를 통해 도출된 가중치 벡터이고, 도 3b는 본 출원의 실험예에 따라 태양광 발전율 예측값들과 실제 발전율 값을 비교한 3일간의 예시 그래프이다. 도 3b에서 빨간색은 본 출원의 실시예를 통해 도출한 발전율 예측값, 연한 초록색은 다른 지역 모델들의 발전율 예측값, 파란색은 실제 발전율 값을 나타낸다.

[0029] 도 2의 예시처럼, 태양광 발전소가 위치한 지역을 훈련 지역으로 삼았다.

[0030] 예측 성능평가 지표로는 시계열 예측에서 가장 많이 활용되는 MAE(Mean Absolute Error)와 MSE(Mean Squared Error)를 활용하였다. MAE와 MSE를 수식으로 표현하면, MAE = $\frac{\sum_{t=1}^{n}|Y_t-\hat{Y}_t|}{n}$,MSE = $\frac{\sum_{t=1}^{n}(Y_t-\hat{Y}_t)^2}{n}$ 와 같다. 수식에서 Y_t는 실제 발전율 값을, $\hat{Y}_t$는 예측된 발전율 값을 뜻한다.

[0031] 실험 결과는 테스트 지역에 대해 제안하는 기법으로 도출한 발전율 예측값과 각 지역별로 학습한 모델에 테스트 지역의 입력 변수를 사용한 발전율 예측값을 테스트 지역의 실제 발전율 값과 각 평가 지표로 비교하였다. 7개 지역을 각각 테스트 지역으로 하여 평가 지표별로 비교한 결과표는 표 1, 2와 같다.

표 1

[0032] MAE 평가 지표

MAE		목표 지역						
		진주	제주	부산	영월	인천	하동	신안
훈련 지역	진주	-	0.345	0.464	0.811	0.443	0.337	0.330
	제주	0.384	-	0.432	0.336	0.412	0.523	0.346
	부산	0.293	0.304	-	0.445	0.338	0.414	0.289
	영월	0.426	0.451	0.492	-	0.584	0.470	0.451
	인천	0.382	0.391	0.344	0.427	-	0.382	0.347
	하동	0.244	0.330	0.382	0.345	0.379	-	0.318
	신안	0.416	0.322	0.467	**0.269**	0.546	0.544	-
제안 기법		**0.235**	**0.281**	**0.332**	0.371	**0.312**	**0.324**	**0.261**

표 2

[0033] MSE 평가지표

MSE	목표 지역						
	진주	제주	부산	영월	인천	하동	신안

훈련 지역	진주	-	0.444	0.589	1.627	0.528	0.398	0.370
	제주	0.535	-	0.531	0.374	0.48	0.953	0.430
	부산	0.311	0.332	-	0.626	0.413	0.585	0.314
	영월	0.602	0.735	0.759	-	0.979	0.762	0.717
	인천	0.473	0.434	0.371	0.587	-	0.474	0.434
	하동	0.230	0.412	0.479	0.472	0.420	-	0.394
	신안	0.639	0.383	0.768	**0.273**	1.098	1.103	-
제안 기법		**0.186**	**0.265**	**0.334**	0.319	**0.345**	**0.329**	**0.234**

[0034] 테스트 지역 한 곳에 대해 지역별로 학습된 모델로 결과를 예측한 평가 지표 값은 같은 열에 각각 표시된다. 예를 들어, 표 2에서 진주 지역의 기상, 발전율 데이터로 학습된 모델에 제주 지역의 기상 데이터를 입력하여 예측한 발전율 값의 MAE 평가 지표값은 0.345이다. 굵은 글씨로 표시된 지표는 테스트 지역에서 가장 높은 성능을 보인 경우이다. 표 2와 3에서 확인할 수 있듯이 영월 지역을 제외하고 모든 지역에서 제안하는 기법이 가장 우수한 성능을 보였다.

[0035] 도 3a를 참조하면, 진주 지역은 인천, 신안, 제주 지역에 비해 하동, 부산, 영월 지역과 더 높은 가중치 값을 가지는 것을 확인할 수 있다.

[0036] 도 3b를 참조하면, 다른 지역에서 훈련한 모델들의 발전율 예측값들은 실제 발전율 값과 큰 편차를 보이는 경우가 많지만, 본 출원의 실시예를 통한 발전율 예측값은 실제 발전율 값과 가장 유사한 형태를 보이는 것을 확인할 수 있다.

[0037] 도 4는 본 출원의 일 실시예에 따른 태양광 발전율 예측 장치의 동작 방법을 보여주는 순서도이다.

[0038] S110 단계에서, 기상관측 데이터와 대기질 특정 데이터를 시계열적으로 수집할 수 있다. 구체적으로, 태양광 발전율 예측 모델을 구성하기 위해 발전량 정보가 풍부한 여러 지역으로부터 기상 정보와 태양광 발전량 정보 등을 수집할 수 있다. 예를 들어, 공공데이터 포털, 기상자료 개방 포털, 에어 코리아 세 곳에서 제공하는 데이터를 사용할 수 있다.

[0039] S120 단계에서, 수집한 데이터에 대해서, 표준 정규화를 수행할 수 있다. 구체적으로, 발전소 별로 설비용량이 다른점을 보완하기 위해 발전량을 설비용량으로 나눈 발전율 값을 출력 변수로 활용하고, 수집한 발전율 데이터, 기상관측 데이터, 대기질 측정 데이터의 범위가 각각 상이하므로 평균은 0으로, 표준 편차를 1로 정규화하는 표준 정규화를 수행한다.

[0040] S130 단계에서, 테스트 지역 예측에 필요한 딥러닝 모델을 각 훈련 지역마다 학습할 수 있다. 구체적으로, 전처리된 기상관측 데이터, 대기질 측정 데이터는 입력변수로 활용하고, 발전율 데이터는 출력변수로 활용하여 딥러닝 모델을 학습한다.

[0041] S140 단계에서, 상기 각 훈련 지역에서 학습된 상기 딥러닝 모델에 상기 테스트 지역 입력 변수를 활용하여 발전율 예측값 벡터를 생성할 수 있다.

[0042] S150 단계에서, 상기 훈련 지역과 상기 테스트 지역의 기상 유사도를 기반으로 가중치를 설정할 수 있다. 구체적으로, 각 훈련 지역에서 입력변수로 사용된 기상관측 데이터, 대기질 측정 데이터와 출력 변수로 사용된 발전율 데이터간의 상관관계를 구한 행렬과 각 훈련 지역의 기상관측 데이터, 대기질 측정 데이터와 테스트 지역의 기상관측 데이터, 대기질 측정 데이터 간의 상관관계를 구한 행렬을 지역별로 내적값을 도출하여 이어붙인 벡터를 생성하고 벡터 값의 합이 1이 되도록 조정하여 최종 가중치 벡터를 생성한다.

[0043] S160 단계에서, 상기 발전율 예측값 벡터와 상기 가중치 벡터를 활용하여 상기 테스트 지역의 발전율을 예측할 수 있다. 구체적으로, 테스트 지역 발전율 예측부(400)에서 생성한 발전율 예측값 벡터와 가중치 벡터 생성부(500)에서 생성한 가중치 벡터를 내적곱하여 테스트 지역에 대한 최종적인 발전율 예측값을 도출한다.

[0044] 도 5는 본 출원의 일 실시예에 따라 기상 유사도를 통하여 가중치를 구하는 과정을 나타내는 순서도이다.

[0045] S210 단계에서, 먼저 i번째 지역에 대해서 발전율과 V개 변수 간의 상관관계를 도출한 크기 $1\times V$의 벡터를 생성한다. 해당 벡터는 상관관계 행렬 $Corr(Y_{train}, X_{train})$의 i번째 열에 해당하며, 변수는 기상관측 데이터 및 대기질 측정 데이터에 관한 것이다.

[0046] S220 단계에서, i번째 지역과 테스트 지역의 각 기상 변수의 피어슨 상관관계를 도출한 크기 $1\times V$의 벡터를 생성한다. 해당 벡터는 상관관계 행렬 $Corr(X_{train}, X_{test})$의 i번째 열에 해당하며, 각 기상 변수는 변수는 기상관측 데이터 및 대기질 측정 데이터에 관한 것이다.

[0047] S230 단계에서, 앞에서 생성된 두 벡터를 내적하여 나온 크기 $1\times N$의 벡터에 대하여 모든 요소의 합이 1이 되도록 조정하면 W_{test}'이 도출된다.

[0048] 이상에서는 본 발명에 따른 바람직한 실시 예들에 대하여 도시하고 또한 설명하였다. 그러나 본 발명은 상술한 실시 예에 한정되지 아니하며, 특허 청구의 범위에서 첨부하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 누구든지 다양한 변형 실시가 가능할 것이다.

부호의 설명

[0049] 100: 데이터 수집부

200: 데이터 전처리부

300: 딥러닝 모델 학습부

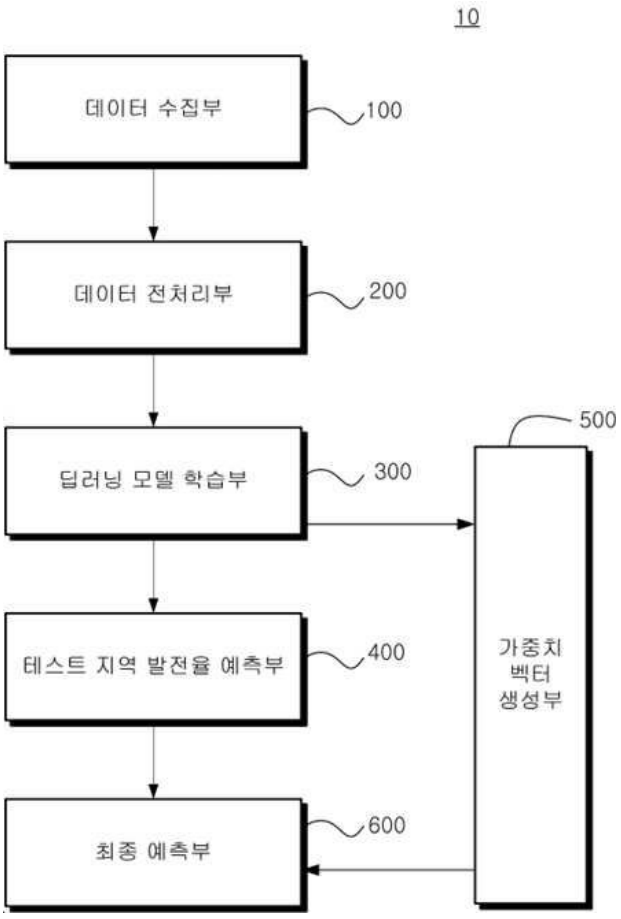
400: 테스트 지역 발전율 예측부

500: 가중치 벡터 생성부

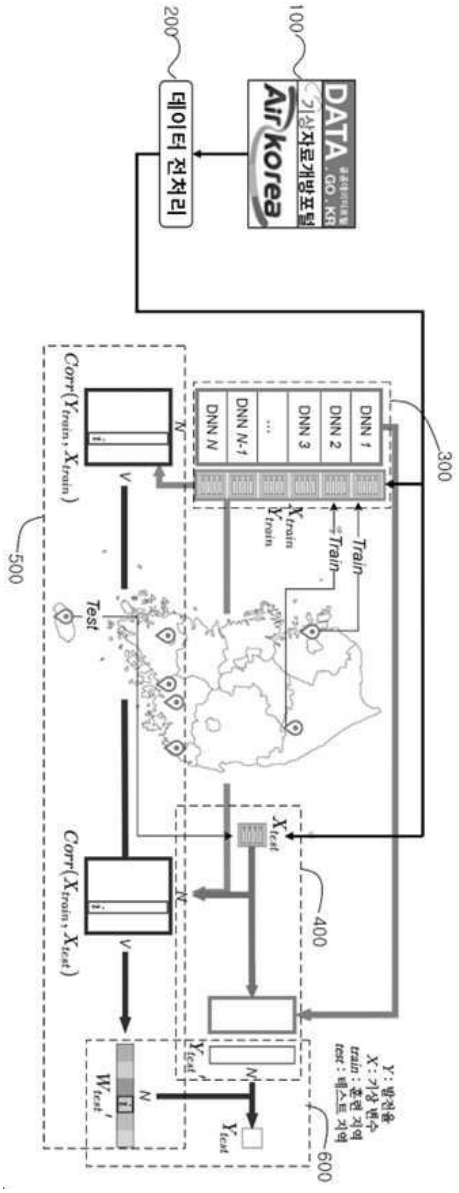
600: 최종 발전율 예측부

도면

도면1



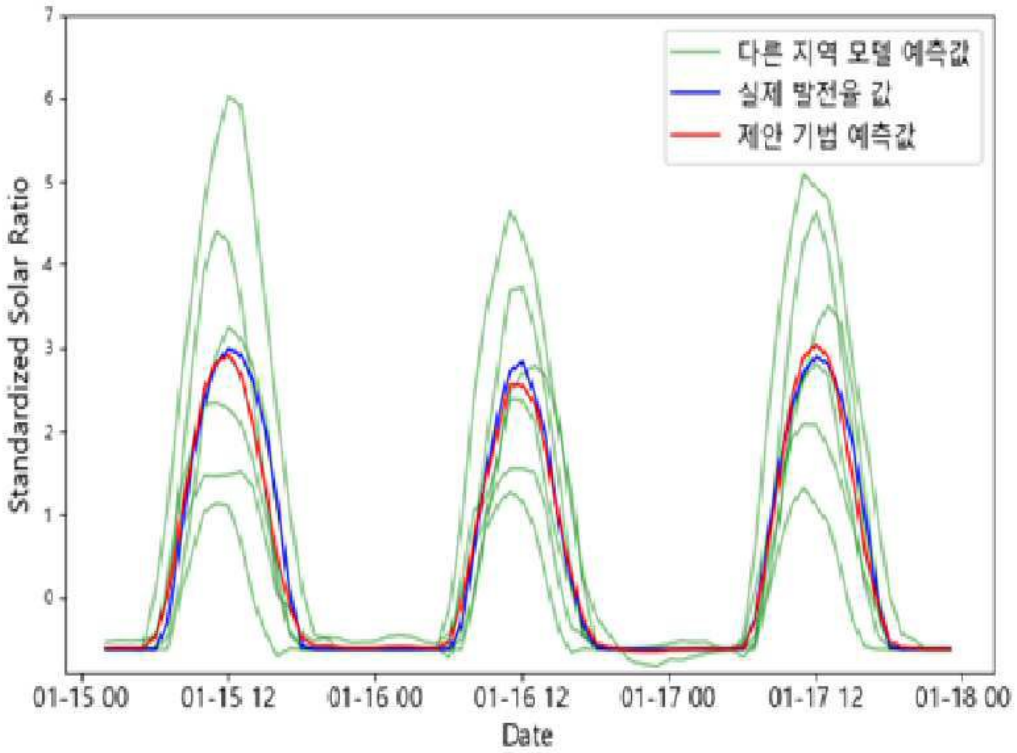
도면2


100
200
300
400
500
600
데이터 전처리
DATA
Airkorea
DNN 1
DNN 2
DNN 3
DNN N-1
DNN N
Train
Test

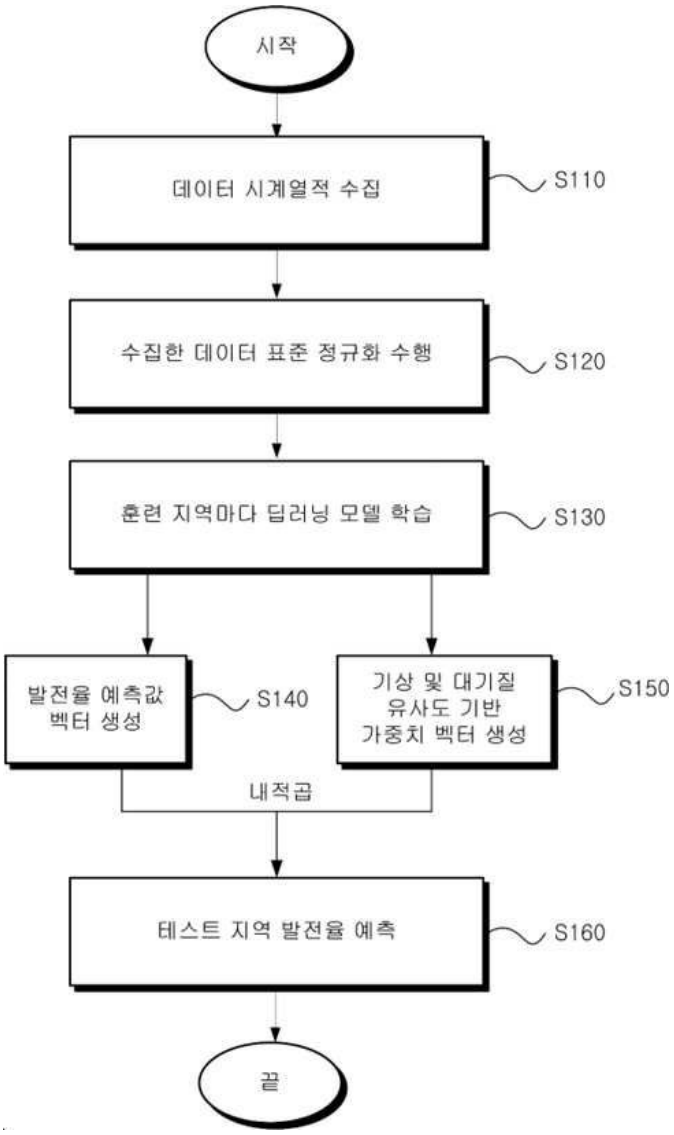
도면3a


가중치
0.13
0.18
0.18
0.15
0.19
0.16
제주
부산
영월
인천
하동
신안
0.20
0.15

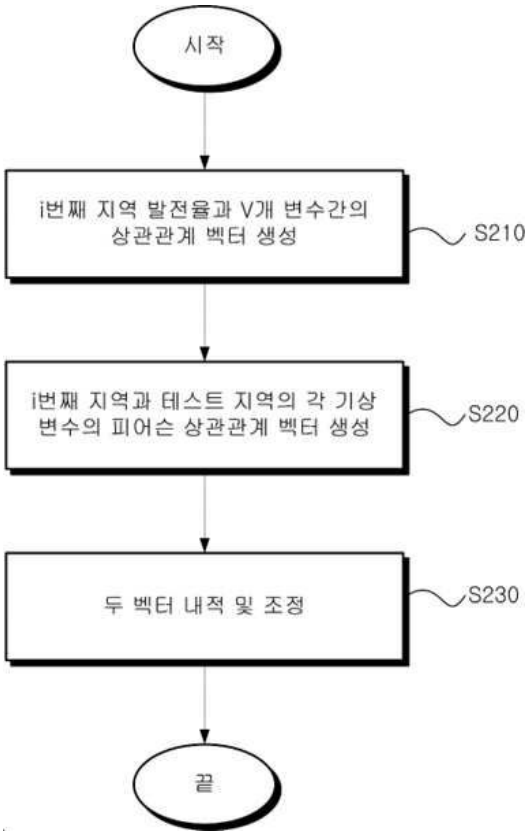
도면3b


다른 지역 모델 예측값
실제 발전율 값
제안 기법 예측값
Standardized Solar Ratio
0
1
2
3
4
5
6
7
01-15 00
01-15 12
01-16 00
01-16 12
01-17 00
01-17 12
01-18 00
Date

도면4


시작
데이터 시계열적 수집
S110
수집한 데이터 표준 정규화 수행
S120
훈련 지역마다 딥러닝 모델 학습
S130
발전율 예측값 벡터 생성
S140
기상 및 대기질 유사도 기반 가중치 벡터 생성
S150
내적곱
테스트 지역 발전율 예측
S160
끝

도면5


시작
i번째 지역 발전율과 V개 변수간의 상관관계 벡터 생성
S210
i번째 지역과 테스트 지역의 각 기상 변수의 피어슨 상관관계 벡터 생성
S220
두 벡터 내적 및 조정
S230
끝