



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년06월03일
(11) 등록번호 10-2405484
(24) 등록일자 2022년05월31일

- (51) 국제특허분류(Int. Cl.)
G06N 3/04 (2006.01) G06N 3/08 (2006.01)
- (52) CPC특허분류
G06N 3/049 (2013.01)
G06N 3/08 (2013.01)
- (21) 출원번호 10-2019-0159705
(22) 출원일자 2019년12월04일
심사청구일자 2019년12월04일
- (65) 공개번호 10-2021-0069898
(43) 공개일자 2021년06월14일
- (56) 선행기술조사문헌
비특허문헌 2(한국데이터정보과학회, 2019.05.31)
비특허문헌 3(한국융합학회, 2018.06.28)
비특허문헌 4(한국디지털정책학회, 2018.05.28)
KR101657495 B1
- (73) 특허권자
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
- (72) 발명자
유성준
서울특별시 광진구 능동로 209, 대양AI센터 719호(군자동)
구영현
서울특별시 광진구 능동로 209, 대양AI센터 717호(군자동)
(뒷면에 계속)
- (74) 대리인
양성보

전체 청구항 수 : 총 8 항

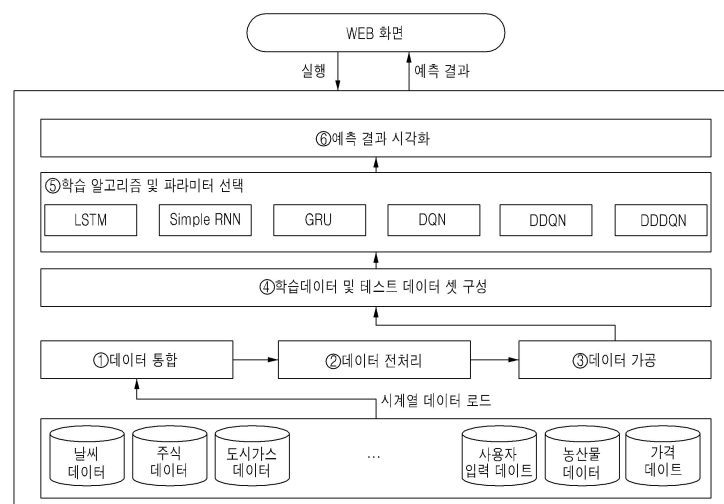
심사관 : 노지명

(54) 발명의 명칭 시계열 분석 간편화를 위한 자동화된 딥러닝 스튜디오

(57) 요약

시계열 분석 간편화를 위한 자동화된 딥러닝 스튜디오가 개시된다. 일 실시예에 따른 딥러닝 시스템에 의해 수행되는 시계열 데이터의 분석을 위한 딥러닝 모델 생성 방법은, 사용자로부터 시계열 데이터가 선택됨을 수신하는 단계; 상기 수신된 시계열 데이터에 기초하여 시계열 데이터의 분석을 위하여 제공된 유저 인터페이스를 통하여 사용자 입력에 따라 딥러닝 모델을 생성하는 단계; 및 상기 생성된 딥러닝 모델을 이용하여 상기 수신된 시계열 데이터를 훈련시킴에 따라 시계열 데이터를 분석한 분석 결과를 제공하는 단계를 포함할 수 있다.

대표도



(72) 발명자

정다운

서울특별시 광진구 능동로 209, 대양AI센터 717호
(군자동)

윤학림

서울특별시 광진구 능동로 209, 대양AI센터 717호
(군자동)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711094394(세부과제번호: 2019-0-00136-001)
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	인공지능융합선도프로젝트
연구과제명	스마트시티 산업 생산성 혁신을 위한 AI융합 기술 개발
기 여 율	1/1
과제수행기관명	세종대학교 산학협력단
연구기간	2019.04.01 ~ 2019.12.31

명세서

청구범위

청구항 1

딥러닝 시스템에 의해 수행되는 시계열 데이터의 분석을 위한 딥러닝 모델 생성 방법에 있어서,

사용자로부터 시계열 데이터가 선택됨을 수신하는 단계;

상기 수신된 시계열 데이터에 기초하여 시계열 데이터의 분석을 위하여 제공된 유저 인터페이스를 통하여 사용자 입력에 따라 딥러닝 모델을 생성하는 단계; 및

상기 생성된 딥러닝 모델을 이용하여 상기 수신된 시계열 데이터를 훈련시킴에 따라 시계열 데이터를 분석한 분석 결과를 제공하는 단계

를 포함하고,

상기 수신하는 단계는,

상기 사용자로부터 입력 또는 선택된 시계열 데이터에 대한 정형 데이터 자동 전처리를 수행하고, 상기 전처리가 수행된 시계열 데이터를 기 설정된 비율로 훈련 데이터 및 테스트 데이터로 분할하고, 상기 분할된 훈련 데이터 중 정답(Target) 데이터를 선택하여 데이터를 분할하고, 상기 선택된 정답 데이터의 분포를 라인 차트(Line Chart) 형식으로 제공하고, 상기 제공된 라인 차트 형식의 정답 데이터의 분포를 통하여 정답 데이터의 이상치가 확인되는 단계

를 포함하는 딥러닝 모델 생성 방법.

청구항 2

제1항에 있어서,

상기 딥러닝 모델을 생성하는 단계는,

상기 수신된 시계열 데이터에 대한 딥러닝 모델을 생성하기 위한 강화학습 알고리즘, LSTM 알고리즘을 포함하는 유저 인터페이스를 제공하고, 상기 제공된 유저 인터페이스를 통하여 학습 알고리즘이 선택됨을 수신하는 단계

를 포함하는 딥러닝 모델 생성 방법.

청구항 3

제2항에 있어서,

상기 딥러닝 모델을 생성하는 단계는,

상기 수신된 학습 알고리즘에 기초하여 하이퍼 파라미터를 설정함에 따라 딥러닝 모델을 구성하는 단계

를 포함하는 딥러닝 모델 생성 방법.

청구항 4

제3항에 있어서,

상기 딥러닝 모델을 생성하는 단계는,

상기 수신된 학습 알고리즘의 에포크 수, 타임 스텝, 배치 사이즈, 레이어 개수, 학습률, 드롭 아웃, 손실 함수, 최적화 함수 또는 활성화 함수를 포함하는 하이퍼 파라미터를 설정하기 위한 인터페이스를 제공하고, 상기 제공된 유저 인터페이스를 통하여 설정된 하이퍼 파라미터를 수신하는 단계

를 포함하는 딥러닝 모델 생성 방법.

청구항 5

제1항에 있어서,
 상기 수신된 시계열 데이터를 기 설정된 비율로 훈련 데이터와 테스트 데이터로 분할하고,
 상기 딥러닝 모델을 생성하는 단계는,
 상기 수신된 시계열 데이터에서 분할된 훈련 데이터를 이용하여 상기 생성된 딥러닝 모델을 훈련시키는 단계
 를 포함하는 딥러닝 모델 생성 방법.

청구항 6

제1항에 있어서,
 상기 수신된 시계열 데이터를 기 설정된 비율로 훈련 데이터와 테스트 데이터로 분할하고,
 상기 분석 결과를 제공하는 단계는,
 상기 수신된 시계열 데이터에서 분할된 테스트 데이터가 상기 수신된 시계열 데이터에서 분할된 분할 데이터를
 이용하여 훈련시킨 딥러닝 모델에 입력됨에 따라 딥러닝 모델의 테스트 결과를 획득하는 단계
 를 포함하는 딥러닝 모델 생성 방법.

청구항 7

제6항에 있어서,
 상기 분석 결과를 제공하는 단계는,
 상기 획득된 테스트 결과의 실측치와 예측치를 라인 차트 형식으로 제공하는 단계
 를 포함하고,
 상기 테스트 결과는 정답 데이터, 예측 기간, 모델 정확도 및 다음달 예측을 포함하는 것을 특징으로 하는 딥러
 닥 모델 생성 방법.

청구항 8

삭제

청구항 9

삭제

청구항 10

딥러닝 시스템에 있어서,
 사용자로부터 시계열 데이터가 선택됨을 수신하는 수신부;
 상기 수신된 시계열 데이터에 기초하여 시계열 데이터의 분석을 위하여 제공된 유저 인터페이스를 통하여 사용
 자 입력에 따라 딥러닝 모델을 생성하는 생성부; 및
 상기 생성된 딥러닝 모델을 이용하여 상기 수신된 시계열 데이터를 훈련시킴에 따라 시계열 데이터를 분석한 분
 석 결과를 제공하는 제공부
 를 포함하고,
 상기 수신부는,
 상기 사용자로부터 입력 또는 선택된 시계열 데이터에 대한 정형 데이터 자동 전처리를 수행하고, 상기 전처리
 가 수행된 시계열 데이터를 기 설정된 비율로 훈련 데이터 및 테스트 데이터로 분할하고, 상기 분할된 훈련 데
 이터 중 정답(Target) 데이터를 선택하여 데이터를 분할하고, 상기 선택된 정답 데이터의 분포를 라인 차트
 (Line Chart) 형식으로 제공하고, 상기 제공된 라인 차트 형식의 정답 데이터의 분포를 통하여 정답 데이터의
 이상치가 확인되는

딥러닝 시스템.

발명의 설명

기술 분야

[0001] 아래의 설명은 시계열 데이터의 분석을 위한 딥러닝 모델을 생성하는 기술에 관한 것이다.

배경 기술

[0003] 가격 예측, 주식 예측, 온도 예측 등 시계열 데이터 분석에 관해서 많은 연구가 이루어지고 있다. 기존 예측 시스템들은 통계적 기반 분석, 머신러닝, 딥러닝 AI 등을 사용자가 코딩을 통해 구현하고 예측하는 시스템이다. 방대한 데이터를 활용해 AI를 만들어 분석가들이 사용하고 있지만 비전문가들은 프로그래밍 언어와 인공지능 분야 지식이 없기 때문에 개발하는데 어려움을 겪고 있다. 또한 분석가들도 자동화 되어있지 않기 때문에 불편함을 겪고 있다.

[0004] 한편, 한국공개특허 제10-2019-0114694호에 시계열 데이터 분석 기술에 관한 것으로, 시계열 데이터가 시간 축에서 스플릿된 복수개의 유닛 각각에 대하여, 유닛 각각의 피쳐(feature)를 중간 인공신경망에 입력하고, 중간 인공신경망으로부터 m (m 은 2 이상의 자연수) 차원의 중간 출력 데이터를 얻고, 시간적으로 바로 인접한 복수개의 유닛의 중간 출력 데이터를 최종 인공신경망에 입력하여 최종 인공신경망으로부터 출력된 최종 출력 데이터를 얻고, 최종 출력 데이터를 이용하여 시계열 데이터의 분석 결과를 생성하는 구성이 개시된 바 있다.

발명의 내용

해결하려는 과제

[0006] 비전문가 및 분석가들이 프로그래밍 언어와 인공지능 분야에 지식이 없어도 손쉽게 개인이 가지고 있는 지식을 활용해 수치를 예측할 수 있는 자동화된 딥러닝 스튜디오(Automated Deep Learning Studio)를 제공하는 방법 및 시스템을 제안할 수 있다.

과제의 해결 수단

[0008] 딥러닝 시스템에 의해 수행되는 시계열 데이터의 분석을 위한 딥러닝 모델 생성 방법은, 사용자로부터 시계열 데이터가 선택됨을 수신하는 단계; 상기 수신된 시계열 데이터에 기초하여 시계열 데이터의 분석을 위하여 제공된 유저 인터페이스를 통하여 사용자 입력에 따라 딥러닝 모델을 생성하는 단계; 및 상기 생성된 딥러닝 모델을 이용하여 상기 수신된 시계열 데이터를 훈련시킴에 따라 시계열 데이터를 분석한 분석 결과를 제공하는 단계를 포함할 수 있다.

[0009] 상기 딥러닝 모델을 생성하는 단계는, 상기 수신된 시계열 데이터에 대한 딥러닝 모델을 생성하기 위한 강화학습 알고리즘, LSTM 알고리즘을 포함하는 유저 인터페이스를 제공하고, 상기 제공된 유저 인터페이스를 통하여 학습 알고리즘이 선택됨을 수신하는 단계를 포함할 수 있다.

[0010] 상기 딥러닝 모델을 생성하는 단계는, 상기 수신된 학습 알고리즘에 기초하여 하이퍼 파라미터를 설정함에 따라 딥러닝 모델을 구성하는 단계를 포함할 수 있다.

[0011] 상기 딥러닝 모델을 생성하는 단계는, 상기 수신된 학습 알고리즘의 에포크 수, 타임 스텝, 배치 사이즈, 레이어 개수, 학습률, 드롭 아웃, 손실 함수, 최적화 함수 또는 활성화 함수를 포함하는 하이퍼 파라미터를 설정하기 위한 인터페이스를 제공하고, 상기 제공된 유저 인터페이스를 통하여 설정된 하이퍼 파라미터를 수신하는 단계를 포함할 수 있다.

[0012] 상기 딥러닝 모델을 생성하는 단계는, 상기 수신된 시계열 데이터에서 사용되는 훈련 데이터에 분할된 학습 데이터를 상기 생성된 딥러닝 모델을 통하여 훈련시키는 단계를 포함할 수 있다.

[0013] 상기 분석 결과를 제공하는 단계는, 상기 수신된 시계열 데이터에서 사용되는 훈련 데이터에 분할된 테스트 데이터를 상기 수신된 시계열 데이터에서 사용되는 훈련 데이터에 분할된 학습 데이터를 훈련시킨 딥러닝 모델에 훈련시킴에 따라 딥러닝 모델의 예측 결과를 획득하는 단계를 포함할 수 있다.

- [0014] 상기 분석 결과를 제공하는 단계는, 상기 획득된 예측 결과의 실측치와 예측치를 라인 차트 형식으로 제공하는 단계를 포함하고, 상기 예측 결과는 정답 데이터, 예측 기간, 모델 정확도 및 다음달 예측을 포함할 수 있다.
- [0015] 상기 수신하는 단계는, 상기 사용자로부터 입력 또는 선택된 시계열 데이터에 대한 정형 데이터 자동 전처리를 수행하고, 상기 전처리가 수행된 시계열 데이터를 훈련 데이터로 사용하고, 상기 사용되는 훈련 데이터를 기 설정된 비율로 학습 데이터 및 테스트 데이터의 비율로 분할하는 단계를 포함할 수 있다.
- [0016] 상기 수신하는 단계는, 상기 사용되는 훈련 데이터 중 정답(Target) 데이터를 선택하여 데이터를 분할하고, 상기 선택된 정답 데이터의 분포를 라인 차트(Line Chart) 형식으로 제공하고, 상기 제공된 라인 차트 형식의 정답 데이터의 분포를 통하여 정답 데이터의 이상치가 확인되는 단계를 포함할 수 있다.
- [0017] 딥러닝 시스템은, 사용자로부터 시계열 데이터가 선택됨을 수신하는 수신부; 상기 수신된 시계열 데이터에 기초하여 시계열 데이터의 분석을 위하여 제공된 유저 인터페이스를 통하여 사용자 입력에 따라 딥러닝 모델을 생성하는 생성부; 및 상기 생성된 딥러닝 모델을 이용하여 상기 수신된 시계열 데이터를 훈련시킴에 따라 시계열 데이터를 분석한 분석 결과를 제공하는 제공부를 포함할 수 있다.

발명의 효과

- [0019] 일 실시예에 따른 딥러닝 시스템은 사용자가 프로그래밍 언어, 인공지능에 대한 지식이 없어도 수치를 예측하는 딥러닝 모델을 간편하게 생성할 수 있다. 또한, 지식이 있는 사람도 프로그래밍을 하지 않고 사용자가 원하는 딥러닝 모델을 생성할 수 있다.
- [0020] 일 실시예에 따른 딥러닝 시스템은 자동화된 딥러닝 스튜디오를 사용하는 사용자들이 큰 지식이 없이 사용자 본인이 원하는 딥러닝 모델을 생성하고, 생성된 딥러닝 모델을 통하여 시계열 데이터의 분석에 다양하게 활용할 수 있도록 제공할 수 있다.

도면의 간단한 설명

- [0022] 도 1은 일 실시예에 따른 딥러닝 시스템의 구조도를 설명하기 위한 도면이다.
- 도 2 및 도 3은 일 실시예에 따른 딥러닝 시스템에서 데이터를 입력하는 것을 설명하기 위한 도면이다.
- 도 3 및 도 4는 일 실시예에 따른 딥러닝 시스템에서 CSV 파일을 선택하는 것을 설명하기 위한 도면이다.
- 도 5는 일 실시예에 따른 딥러닝 시스템에서 입력 데이터를 테이블 형식으로 확인하는 것을 설명하기 위한 도면이다.
- 도 6은 일 실시예에 따른 딥러닝 시스템에서 데이터를 분할하는 것을 설명하기 위한 도면이다.
- 도 7은 일 실시예에 따른 딥러닝 시스템에서 타겟 데이터를 선택하는 것을 설명하기 위한 도면이다.
- 도 8은 일 실시예에 따른 딥러닝 시스템에서 타겟 데이터의 분할을 확인하는 것을 설명하기 위한 도면이다.
- 도 9는 일 실시예에 따른 딥러닝 시스템에서 알고리즘을 선택하는 것을 설명하기 위한 도면이다.
- 도 10은 일 실시예에 따른 딥러닝 시스템에서 하이퍼 파라미터를 설정하는 것을 설명하기 위한 도면이다.
- 도 11 및 도 12는 일 실시예에 따른 딥러닝 시스템에서 모델을 테스트하는 것을 설명하기 위한 도면이다.
- 도 13은 일 실시예에 따른 딥러닝 시스템에서 시계열 데이터의 분석을 위한 딥러닝 모델을 생성하는 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0023] 이하, 실시예를 첨부한 도면을 참조하여 상세히 설명한다.
- [0025] 비전문가 및 분석가들이 프로그래밍 언어와 인공지능 분야에 지식이 없어도 손쉽게 개인이 가지고 있는 지식을 활용해 수치를 예측할 수 있는 자동화된 딥러닝 스튜디오(Automated Deep Learning Studio)를 제공하는 방법 및 시스템을 제안할 수 있다. 이때, 자동화 딥러닝 스튜디오는 시계열 데이터의 분석을 위한 딥러닝 모델을 자동 또는 수동으로 생성할 수 있는 유저 인터페이스를 제공하고, 제공된 유저 인터페이스를 통하여 시계열 데이터와 관련된 각각의 파라미터 정보를 입력함으로써 생성된 딥러닝 모델을 통하여 시계열 데이터의 분석을 수행할 수

있다.

- [0026] 구체적으로, 딥러닝 시스템은 스튜디오 형식으로 사용자가 원하는 시계열 데이터를 불러오고, 불러온 데이터로 제공하는 딥러닝(Deep Learning) 알고리즘을 학습시켜 원하는 모델을 생성할 수 있다. 이러한 전체 과정이 웹 화면 또는 모바일 화면에 표시되어 사용자가 클릭 몇 번으로 손쉽게 딥러닝 모델을 생성할 수 있도록 제공될 수 있다. 또한, 생성된 딥러닝 모델을 테스트할 수 있는 환경을 제공하여 사용자로부터 생성된 딥러닝 모델에 정확도를 확인해볼 수 있도록 제공할 수 있다. 이때, 테스트의 결과는 예측기간, 모델 정확도, 다음 예측 값으로 구성될 수 있다. 딥러닝 시스템은 프로그래밍 언어와 인공지능을 모르는 비전문가와 전문가 모두 쉽게 인공지능을 만들 수 있을 것으로 기대한다.
- [0027] 도 1은 일 실시예에 따른 딥러닝 시스템의 구조도를 설명하기 위한 도면이다.
- [0028] 딥러닝 시스템은 시계열 데이터를 입력 및 선정할 수 있는 UI/UX를 제공할 수 있다. 예를 들면, 시계열 데이터는 날씨 데이터, 주식 데이터, 도시가스 데이터, 사용자 입력 데이터, 농산물 데이터 및 가격 데이터 등을 포함할 수 있다. 사용자는 제공된 UI/UX를 통해 시계열 데이터를 입력 및 선정을 할 수 있다. 딥러닝 시스템은 사용자로부터 입력 또는 선정된 시계열 데이터를 불러와 정형데이터 자동 전처리 기술을 통해 데이터 정제, 데이터 정규화, 데이터 분리, 데이터 통합을 진행할 수 있다.
- [0029] 딥러닝 시스템은 자동 전처리 기술을 수행한 시계열 데이터에 적합한 특징을 추출하고 학습 알고리즘 및 파라미터 선택을 통해 딥러닝 모델에 적용할 수 있다. 딥러닝 시스템은 학습된 딥러닝 모델을 검증 가능하게 시각화하여 사용자에게 제공할 수 있다.
- [0030] 도 2 및 도 3은 일 실시예에 따른 딥러닝 시스템에서 데이터를 입력하는 것을 설명하기 위한 도면이다.
- [0031] 딥러닝 시스템은 자동화된 딥러닝 스튜디오를 통하여 시계열 데이터의 분석을 위한 유저 인터페이스를 제공할 수 있다. 이때, 유저 인터페이스는 시계열 데이터와 관련하여 딥러닝 모델을 생성하기 위한 사용자의 명령을 입력받기 위하여 구성된 것을 의미할 수 있다. 도 2를 참고하면, '수유역 대합실'이라는 시계열 데이터의 입력 화면을 나타낸 것이고, 도 3을 참고하면, '농수산_배추'라는 시계열 데이터의 입력 화면을 나타낸 것이다.
- [0032] 딥러닝 시스템은 자동화된 딥러닝 스튜디오를 통하여 시계열 데이터를 입력하는 유저 인터페이스를 제공할 수 있다. 딥러닝 시스템은 시계열 데이터를 포함하는 CSV 파일 형식의 시계열 데이터를 선택하는 유저 인터페이스를 제공할 수 있다. 예를 들면, 사용자는 자동화된 딥러닝 스튜디오에 제공된 'Browse' 버튼을 클릭하여 시계열 데이터를 포함하는 CSV 파일을 선택할 수 있다. 도 4 및 도 5를 참고하면, 딥러닝 시스템은 자동화된 딥러닝 스튜디오를 통하여 시계열 데이터를 포함하는 CSV 파일 리스트를 제공할 수 있고, 제공된 CSV 파일 리스트로부터 사용자로부터 CSV 파일이 선택됨을 수신할 수 있다. 이때, 시계열 데이터의 입력에는 CSV 파일 형식만 입력이 가능할 수 있다. 도 4를 참고하면, '수유역 대합실'이라는 시계열 데이터, 도 5를 참고하면, '농수산_배추'라는 시계열 데이터가 선택될 수 있다. 이외에도 다양한 종류의 시계열 데이터가 선택될 수 있으며, 동시에 적어도 하나 이상의 시계열 데이터가 선택될 수도 있다.
- [0033] 딥러닝 시스템은 입력된 시계열 데이터를 테이블 형식으로 제공할 수 있다. 사용자는 입력된 시계열 데이터를 테이블 형식으로 확인할 수 있다. 구체적으로, 자동화된 딥러닝 스튜디오에 제공된 '파일 확인' 버튼이 클릭됨에 따라 선택된 CSV 파일의 기 설정된 개수(예를 들면, 10개)의 데이터를 테이블 형식으로 확인할 수 있다. 이때, 사용자로부터 설정된 개수의 데이터가 테이블 형식으로 제공될 수 있다.
- [0034] 딥러닝 시스템은 사용자로부터 선택된 CSV 파일의 데이터를 딥러닝 모델의 훈련용 데이터로 사용할 수 있다. 딥러닝 시스템은 자동화된 딥러닝 스튜디오에 데이터 분할을 위한 유저 인터페이스를 제공할 수 있다. 딥러닝 시스템은 자동화된 딥러닝 스튜디오에 제공된 '다음 단계로' 버튼이 클릭됨에 따라 데이터 분할 화면으로 이동할 수 있다.
- [0035] 도 6은 일 실시예에 따른 딥러닝 시스템에서 데이터를 분할하는 것을 설명하기 위한 도면이다.
- [0036] 딥러닝 시스템은 자동화된 딥러닝 스튜디오를 통하여 데이터 분할 화면을 제공할 수 있다. 딥러닝 시스템은 데이터 분할 화면으로 정답 데이터를 선택하고, 정답 데이터의 분포를 라인 차트(Line Chart) 형식으로 제공할 수 있다.
- [0037] 도 7을 참고하면, 딥러닝 시스템은 자동화된 딥러닝 스튜디오를 통하여 타겟 데이터를 선택하는 유저 인터페이스를 제공할 수 있다. 딥러닝 시스템은 사용자로부터 훈련 데이터 중에서 정답(Target) 데이터를 선택할 수 있다. 예를 들면, 사용자는 'CO2'를 정답 데이터로 선택할 수 있고, 'Cabbage'를 정답 데이터로 선택할 수 있다.

이와 같이, 'CO2' 정답 데이터 이외에도 다양한 정답 데이터를 선택할 수 있다. 지도 학습(Supervised Learning)은 데이터에 대한 정답 데이터가 주어진 상태에서 컴퓨터를 학습시키는 방법이다. 딥러닝 시스템은 자동화된 딥러닝 스튜디오를 통하여 정답 데이터를 선택하기 위한 정답 데이터 리스트를 제공할 수 있고, 제공된 타겟 데이터 리스트로부터 정답 데이터가 선택됨을 수신할 수 있다. 예를 들면, 자동화된 딥러닝 스튜디오를 통하여 훈련 데이터 중 정답 데이터가 선택된 후, '데이터 확인하기' 버튼이 선택될 수 있다. 딥러닝 시스템은 자동화된 딥러닝 스튜디오를 통하여 정답 데이터가 분할되는 것을 확인할 수 있다. 이때, 정답 데이터가 분할되는 것이 그래프로 표시될 수 있다.

[0038] 도 8을 참고하면, 자동화된 딥러닝 스튜디오에서 '데이터 확인하기' 버튼이 클릭됨에 따라 선택한 정답 데이터의 분포를 라인 차트(Line Chart) 형식으로 확인할 수 있다. 딥러닝 시스템은 라인 차트 형식으로 제공된 정답 데이터를 통하여 이상치를 확인할 수 있도록 제공할 수 있다. 사용자는 라인 차트에서 정답 데이터의 이상치를 확인할 수 있다. 딥러닝 시스템은 '이전 단계'의 버튼을 제공하여 데이터 입력 화면으로 이동할 수 있도록 제공할 수 있다. 사용자는 데이터 입력 화면으로 이동하여 시계열 데이터를 재입력할 수도 있다.

[0039] 딥러닝 시스템은 타겟 데이터의 분할을 확인함에 따라 '다음 단계'로 이동하는 유저 인터페이스를 제공할 수 있다. 사용자는 '다음 단계'의 버튼을 클릭함에 따라 알고리즘 선택 화면으로 이동될 수 있다. 딥러닝 시스템은 알고리즘 선택 화면으로 이동할 수 있다. 이때, 딥러닝 시스템은 '다음 단계'의 버튼이 클릭됨에 따라 훈련 데이터를 9(학습 데이터): 1(테스트 데이터)의 비율로 분할할 수 있다.

[0040] 도 9는 일 실시예에 따른 딥러닝 시스템에서 알고리즘을 선택하는 것을 설명하기 위한 도면이다.

[0041] 딥러닝 시스템은 자동화된 딥러닝 스튜디오에 알고리즘 선택 화면을 제공할 수 있다. 딥러닝 시스템은 알고리즘을 선택하기 위한 유저 인터페이스를 제공할 수 있다. 딥러닝 시스템은 자동화된 딥러닝 스튜디오에서 제공하는 알고리즘으로 강화학습(Reinforcement Learning), LSTM(Long Short Term Memory)를 포함하는 알고리즘 그룹을 제공할 수 있고, 알고리즘 내부에 각각의 알고리즘에 속하는 세부 알고리즘을 제공할 수 있다. 강화학습 알고리즘 그룹은 알파고통해 유명해진 알고리즘으로 컴퓨터가 환경, 행동, 보상, 규칙을 통해 에이전트를 학습시키는 방법이다. LSTM 알고리즘 그룹은 시계열 특성을 학습하여 예측을 진행하는 대표적인 딥러닝 알고리즘이다. 일례로, 딥러닝 시스템은 강화학습으로 DQN, DDQN, DDDQN 등을 제공할 수 있고, LSTM으로 LSTM, SimpleRNN, GRU 등을 제공할 수 있다. 사용자는 제시된 알고리즘 중 사용하고자 하는 알고리즘을 선택할 수 있다. 이때, 알고리즘은 추가 및 변경될 수 있다. 딥러닝 시스템은 사용자로부터 선택된 알고리즘으로 이미 분할할 학습 데이터를 훈련시킬 수 있다.

[0042] 딥러닝 시스템은 알고리즘 선택의 이전 과정, 다시 말해서, 데이터 분할 화면으로 이동하기 위한 유저 인터페이스를 제공할 수 있고, 알고리즘 선택의 이후 과정, 다시 말해서, 하이퍼 파라미터를 설정하는 화면으로 이동하기 위한 유저 인터페이스를 제공할 수 있다. 딥러닝 시스템은 '이전 단계'의 버튼이 클릭됨에 따라 데이터 분할 화면으로 이동할 수 있고, '다음 단계'의 버튼이 클릭됨에 따라 하이퍼 파라미터를 설정하는 화면으로 이동할 수 있다. 예를 들면, '이전 단계'의 버튼이 클릭됨에 따라 이동된 데이터 분할 화면을 통하여 정답 데이터를 재선택할 수도 있다.

[0043] 도 10은 일 실시예에 따른 딥러닝 시스템에서 하이퍼 파라미터를 설정하는 것을 설명하기 위한 도면이다.

[0044] 딥러닝 시스템은 자동화된 딥러닝 스튜디오를 통하여 알고리즘의 하이퍼 파라미터를 설정하는 화면을 제공할 수 있다. 이때, 하이퍼 파라미터란 일반적으로 딥러닝에서 어떠한 임의의 모델을 학습시킬 때, 사용자가 직접 튜닝(설정) 해주어야 하는 변수를 의미할 수 있다.

[0045] 딥러닝 시스템은 알고리즘의 하이퍼 파라미터를 설정하는 유저 인터페이스를 제공할 수 있다. 사용자는 제공된 유저 인터페이스를 통하여 하이퍼 파라미터를 입력할 수 있다.

[0046] 딥러닝 시스템은 알고리즘 하이퍼 파라미터 설정의 이전 과정, 다시 말해서, 알고리즘 선택 화면으로 이동하기 위한 유저 인터페이스를 제공할 수 있고, 하이퍼 파라미터 설정의 이후 과정, 다시 말해서, 테스트 시작 화면으로 이동하기 위한 유저 인터페이스를 제공할 수 있다. 딥러닝 시스템은 '이전 단계'의 버튼이 클릭됨에 따라 알고리즘 선택 화면으로 이동할 수 있고, '다음 단계'의 버튼이 클릭됨에 따라 테스트 시작 화면으로 이동할 수 있다. 예를 들면, '이전 단계'의 버튼이 클릭됨에 따라 이동된 알고리즘 선택 화면을 통하여 알고리즘을 재선택할 수도 있다. '테스트 시작' 버튼이 클릭됨에 따라 결과 화면으로 이동될 수 있다.

[0047] 딥러닝 시스템은 '테스트 시작' 버튼이 클릭됨에 따라 설정된 하이퍼 파라미터로 딥러닝 모델을 훈련시킬 수 있

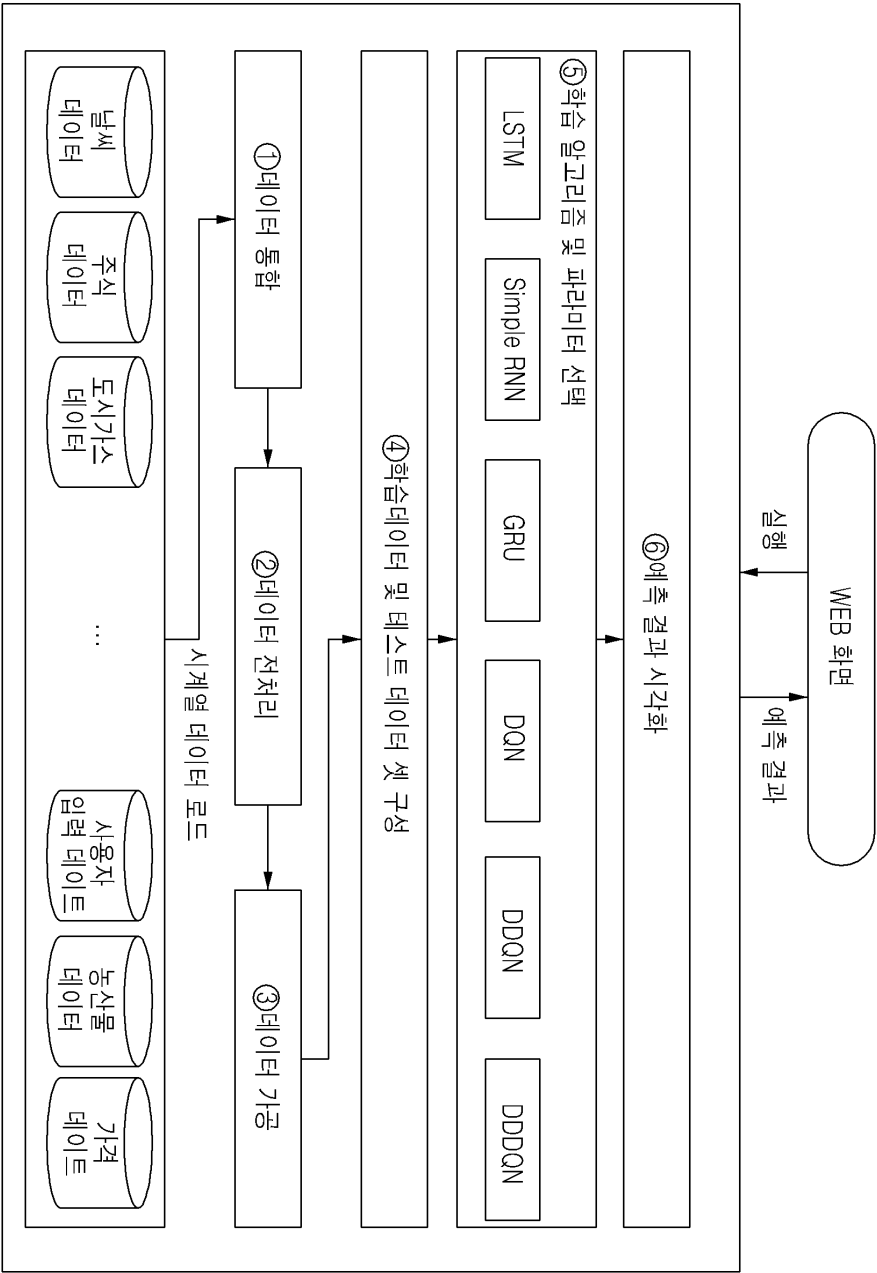
다.

- [0048] 도 11 및 도 12는 일 실시예에 따른 딥러닝 시스템에서 모델을 테스트하는 것을 설명하기 위한 도면이다.
- [0049] 딥러닝 시스템의 자동화된 딥러닝 스튜디오를 통하여 딥러닝 모델의 테스트 결과를 제공할 수 있다. 딥러닝 시스템은 이미 훈련된 딥러닝 모델로 테스트 데이터를 테스트 하여 딥러닝 모델의 예측 결과를 테이블 형식으로 제공할 수 있다. 딥러닝 시스템은 정답 데이터, 예측 기간, 모델 정확도 및 다음달 예측을 포함하는 예측 결과를 제공할 수 있다.
- [0050] 딥러닝 시스템은 테스트 결과의 실측치와 예측치를 라인 차트 형식으로 제공할 수 있다. 사용자는 테스트 결과의 실측치와 예측치를 라인 차트 형식으로 확인할 수 있다. 실측치와 예측치의 분포도 확인할 수 있다.
- [0051] 딥러닝 시스템은 새로운 모델을 생성하기 위한 유저 인터페이스를 제공할 수 있다. 딥러닝 시스템은 '새로운 모델 생성' 버튼이 클릭됨에 따라 데이터 입력 화면으로 이동할 수 있다. 이때, 처음부터 딥러닝 모델을 다시 훈련시킬 수 있다.
- [0052] 도 13은 일 실시예에 따른 딥러닝 시스템에서 시계열 데이터의 분석을 위한 딥러닝 모델을 생성하는 방법을 설명하기 위한 흐름도이다.
- [0053] 단계(1310)에서 딥러닝 시스템은 사용자로부터 시계열 데이터가 선택됨을 수신할 수 있다. 딥러닝 시스템은 사용자로부터 입력 또는 선택된 시계열 데이터에 대한 정형 데이터 자동 전처리를 수행할 수 있다. 딥러닝 시스템은 전처리가 수행된 시계열 데이터를 훈련 데이터로 사용하고, 사용되는 훈련 데이터를 기 설정된 비율로 학습 데이터 및 테스트 데이터의 비율로 분할할 수 있다. 딥러닝 시스템은 사용되는 훈련 데이터 중 정답(Target) 데이터를 선택하여 데이터를 분할하고, 선택된 정답 데이터의 분포를 라인 차트(Line Chart) 형식으로 제공할 수 있다. 이에, 제공된 라인 차트 형식의 정답 데이터의 분포를 통하여 정답 데이터의 이상치가 확인될 수 있다.
- [0054] 단계(1320)에서 딥러닝 시스템은 수신된 시계열 데이터에 기초하여 시계열 데이터의 분석을 위하여 제공된 유저 인터페이스를 통하여 사용자 입력에 따라 딥러닝 모델을 생성할 수 있다. 딥러닝 시스템은 수신된 시계열 데이터에 대한 딥러닝 모델을 생성하기 위한 강화학습 알고리즘, LSTM 알고리즘을 포함하는 유저 인터페이스를 제공하고, 제공된 유저 인터페이스를 통하여 학습 알고리즘이 선택됨을 수신할 수 있다. 딥러닝 시스템은 수신된 학습 알고리즘에 기초하여 하이퍼 파라미터를 설정함에 따라 딥러닝 모델을 구성할 수 있다. 딥러닝 시스템은 수신된 학습 알고리즘의 에포크 수, 타임 스텝, 배치 사이즈, 레이어 개수, 학습률, 드롭 아웃, 손실 함수, 최적화 함수 또는 활성화 함수를 포함하는 하이퍼 파라미터를 설정하기 위한 인터페이스를 제공하고, 제공된 유저 인터페이스를 통하여 설정된 하이퍼 파라미터를 수신할 수 있다. 딥러닝 시스템은 수신된 시계열 데이터에서 사용되는 훈련 데이터에 분할된 학습 데이터를 생성된 딥러닝 모델을 통하여 훈련시킬 수 있다.
- [0055] 단계(1330)에서 딥러닝 시스템은 생성된 딥러닝 모델을 이용하여 수신된 시계열 데이터를 훈련시킴에 따라 시계열 데이터를 분석한 분석 결과를 제공할 수 있다. 딥러닝 시스템은 수신된 시계열 데이터에서 사용되는 훈련 데이터에 분할된 테스트 데이터를 수신된 시계열 데이터에서 사용되는 훈련 데이터에 분할된 학습 데이터를 훈련시킨 딥러닝 모델에 훈련시킴에 따라 딥러닝 모델의 예측 결과를 획득할 수 있다. 딥러닝 시스템은 획득된 예측 결과의 실측치와 예측치를 라인 차트 형식으로 제공할 수 있다. 딥러닝 시스템은 예측 결과로 정답 데이터, 예측 기간, 모델 정확도 및 다음달 예측을 포함할 수 있다.
- [0056] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

- [0057] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.
- [0058] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.
- [0059] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.
- [0060] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

도면1



① 데이터 설정 ② 데이터 분할 ③ 알고리즘 선택 ④ 하이퍼파라미터 설정 ⑤ 결과 확인

데이터 입력

CSV 파일 입력

② 파일 확인 수유역 대합실_235csv ① Browse

③ 날짜 PM10 PM2.5 HCHO CO2 VOC TEMPERATURE HUMIDITY

2019-04-19-00	16	12	0	443	49	23.4	39
2019-04-19-01	16	12	0	440	54	23.4	39
2019-04-19-02	16	12	0	434	51	23.4	39
2019-04-19-03	16	12	0	436	54	23.4	39
2019-04-19-04	21	12	0	432	55	23.4	39
2019-04-19-05	21	15	0	438	55	23.4	39
2019-04-19-06	21	15	0	420	48	23.4	39
2019-04-19-07	16	12	0	435	51	23.4	39
2019-04-19-08	21	15	0	439	52	23.4	39
2019-04-19-09	21	15	0	447	48	23.4	39

④ 다음 단계로

- ① 데이터 설정
- ② 데이터 분할
- ③ 알고리즘 선택
- ④ 하이퍼파라미터 설정
- ⑤ 결과 확인

데이터 입력

CSV 파일 입력

파일 확인

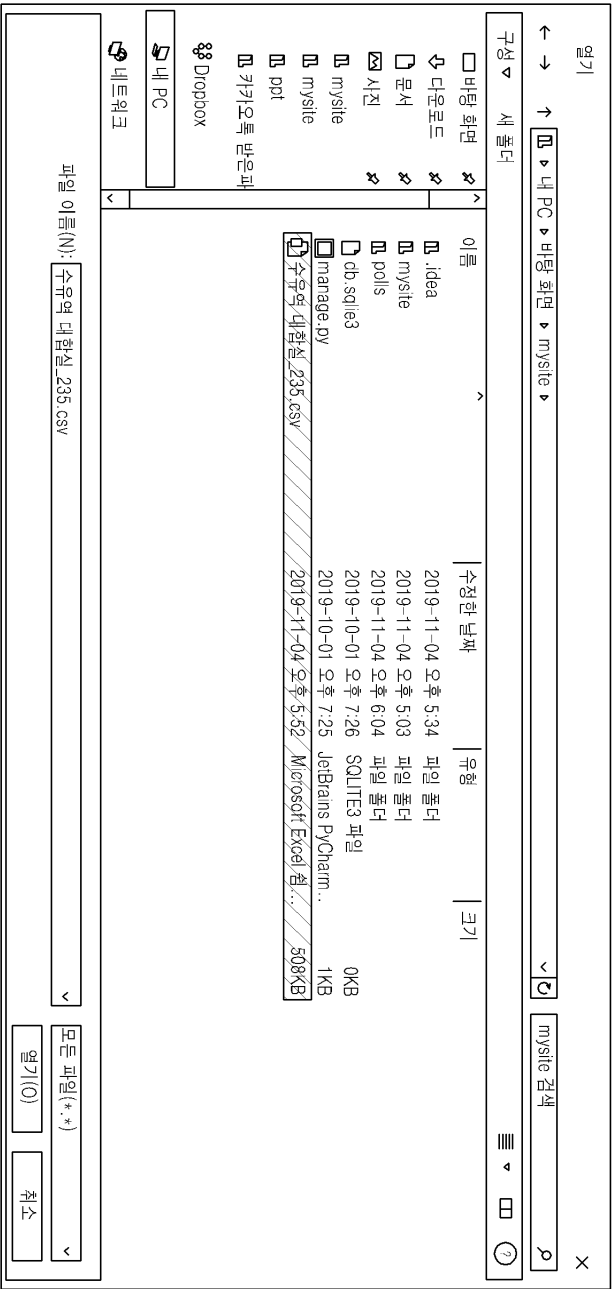
농산물_배추_데이터.csv

Browse

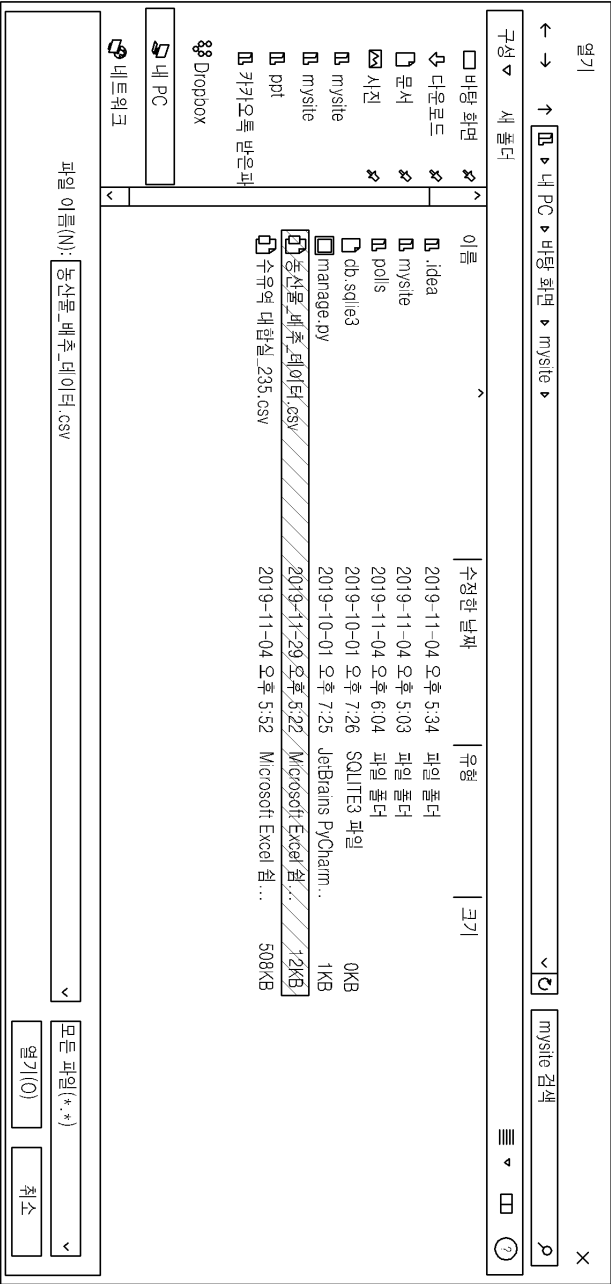
date	cabbage	rp	avg_t	max_t	min_t	humid	rain	sun	month	cp_r
201201	2750	51.59	-0.09	4.13	30.2	47.38	1.22	6.34	1	-2521
201202	4297	7973	0.4	5.09	24.44	41.82	0.79	7.18	2	-1775
201203	7764	8709	5.67	9.55	35.29	56.09	4.08	5.81	3	1142
201204	10179	9084	13.25	18.27	30.07	52.75	4.21	7.03	4	3080
201205	4914	10636	17.58	21.9	41.66	61.96	1.69	7.07	5	-673
201206	4220	11907	20.58	24.1	54.88	74.53	0.86	4.03	6	-518
201207	5589	13071	25.32	28.89	52.94	70.91	8.4	5.89	7	-933
201208	6915	10087	25.03	28.4	63.04	80.63	9.48	2.95	8	-3311
201209	10284	13812	20.18	24.09	49.92	71.42	9.68	5.89	9	-811
201210	7957	17071	15.77	20.8	32.95	55.64	0.33	7.49	10	1381

다음 단계로

도면4



도면5



- ① 데이터 설정 ② 데이터 분할 ③ 알고리즘 선택 ④ 하이퍼파라미터 설정 ⑤ 결과 확인

Target 분할

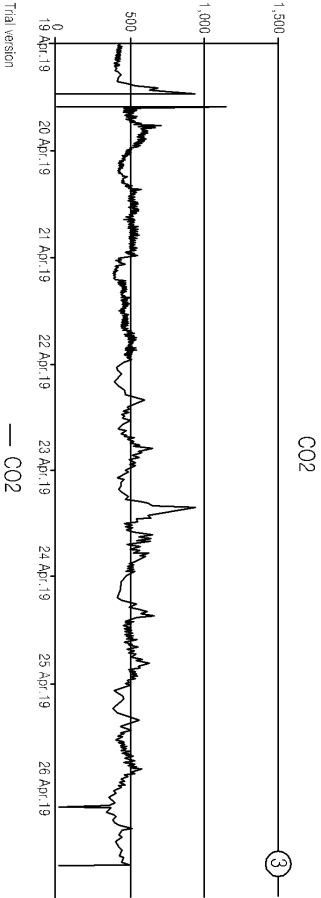
Target 선택

CO2

①

데이터 확인하기

②



- ④ 이전 단계로 ⑤ 다음 단계로

도면6

Target 선택

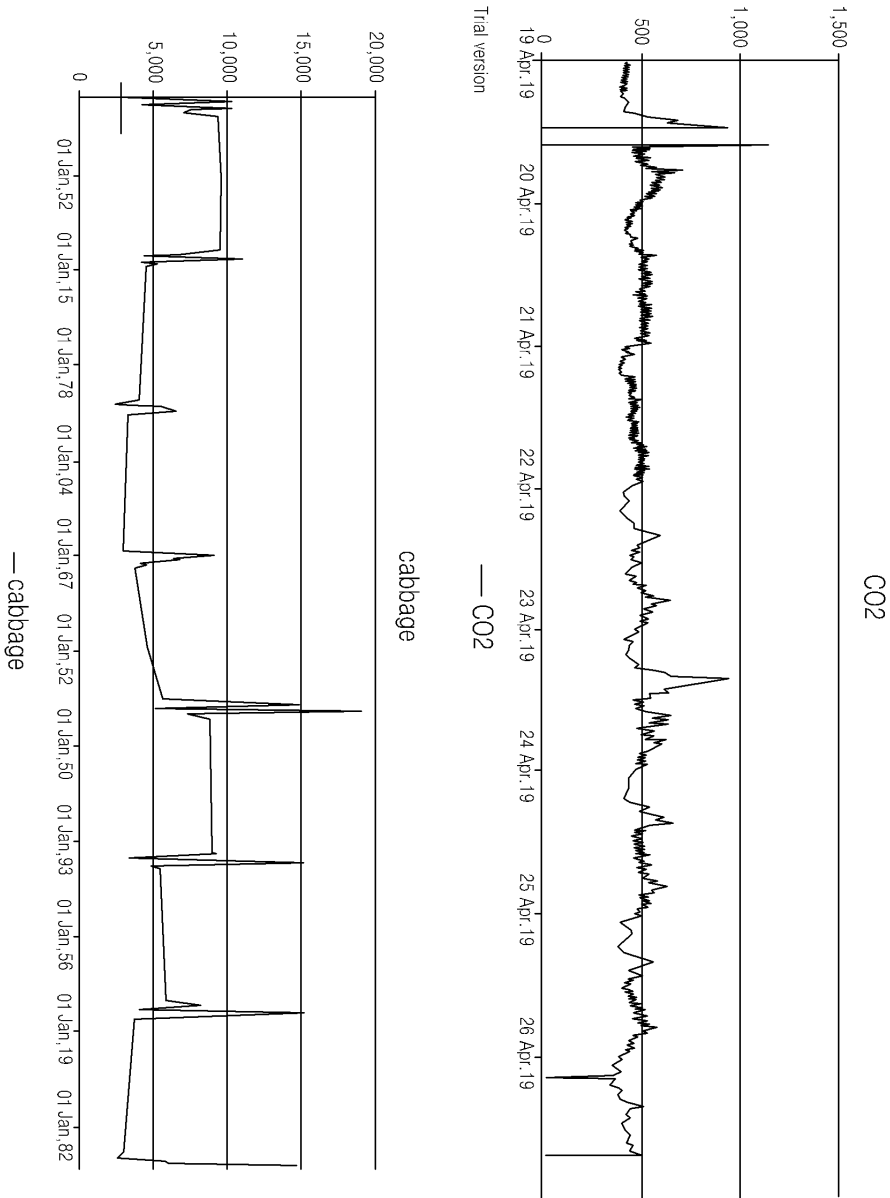
CO2	CO2	▼
	PM10	
데이터	PM2.5	
	HCHO	
	CO2	
	VOC	
	TEMPERATURE	
	HUMIDITY	

Target 분할
CSV 파일 입력

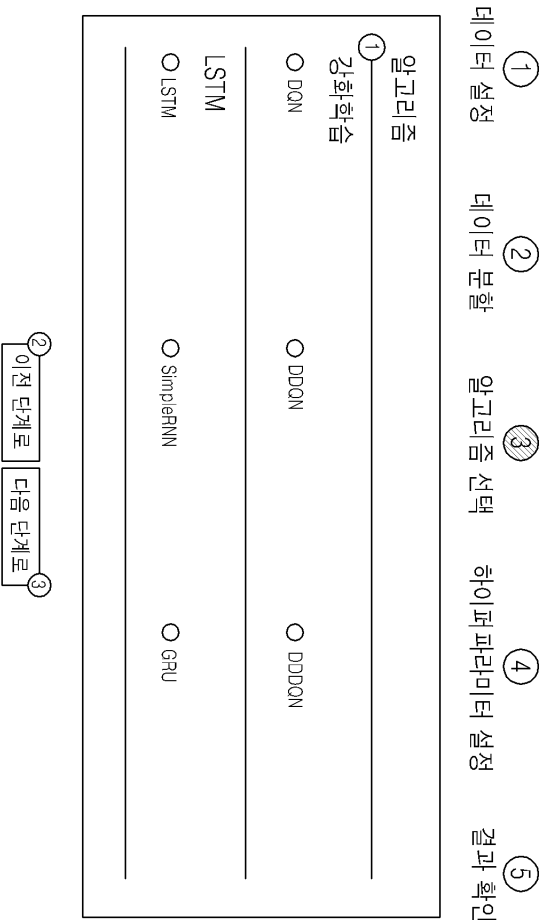
Target 선택	cabbage	▼
	cabbage	
데이터	rp	
	avg_t	
	max_t	
	min_t	
	humid	
	rain	
	sun	
	moth	
	cp_r	

도면7

도면8



도면9



- ① 데이터 설정
- ② 데이터 분할
- ③ 알고리즘 선택
- ④ 하이퍼파라미터 설정
- ⑤ 결과 확인

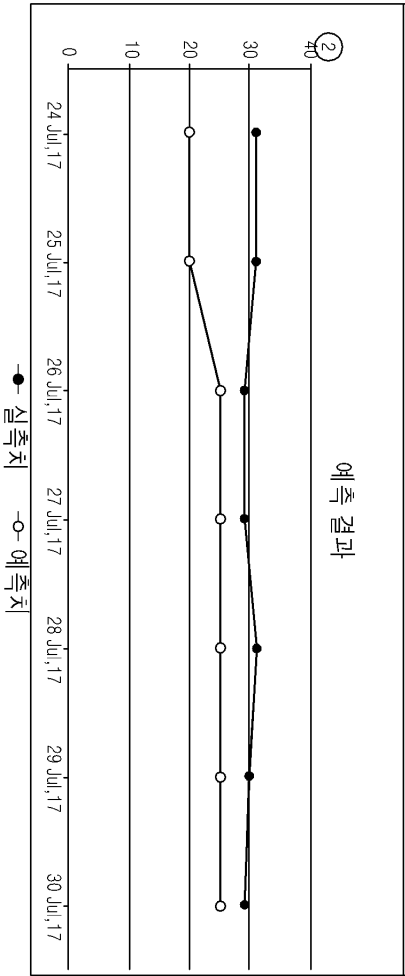
① LSTM

에포크 수	타임 스텝	배치 사이즈
10	4	16
레이어 개수	학습률	드롭아웃
4	0.01	0.3
손실 함수		
mean_squared_error		▽
최적화 함수		
Adam ▽		
활성화 함수		
tanh ▽		

- ② 이전 단계로
- ③ 다음 단계로

- ① 데이터 설정
- ② 데이터 분할
- ③ 알고리즘 선택
- ④ 하이퍼파라미터 설정
- ⑤ 결과 확인

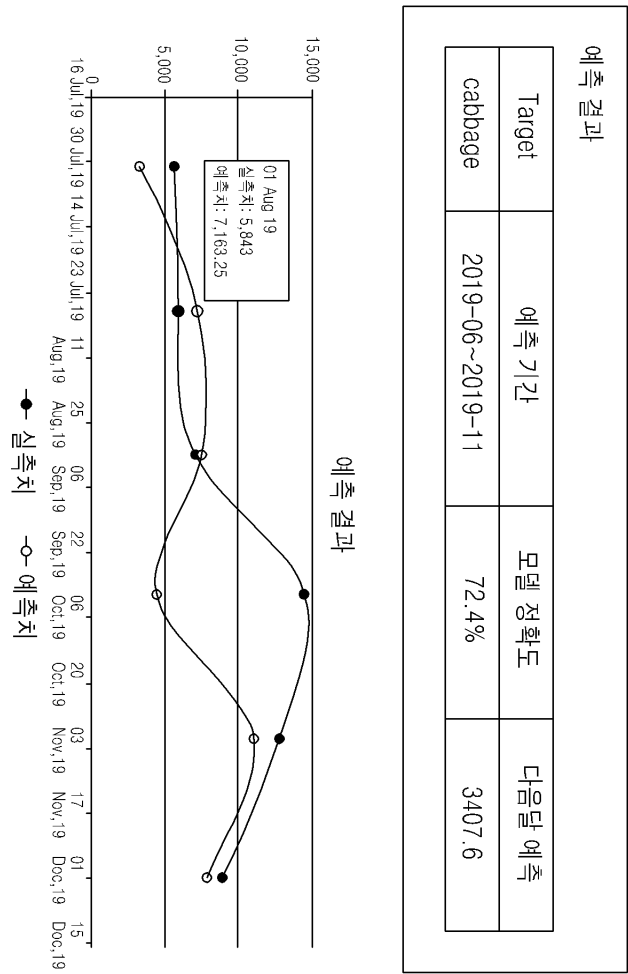
예측 결과			
① Target	예측기간	모델 정확도	다음달 예측
CO2	2019-10-2019-11	85%	460



③ 새로운 모델 생성

도면11

도면12



도면13

