

## 이상저온 발생 시점 확인을 위한 알고리즘 패턴 개발

이정원<sup>1</sup>, 이충호<sup>2\*</sup>

<sup>1</sup>한밭대학교 정보통신공학과 박사, <sup>2</sup>한밭대학교 정보통신공학과 교수

### Development of Algorithm Patterns for Identifying the Time of Abnormal Low Temperature Generation

Jeongwon Lee<sup>1</sup>, Choong Ho Lee<sup>2\*</sup>

<sup>1</sup>Doctoral Student, Dept. of Information and Communication Engineering, Hanbat National University

<sup>2</sup>Professor, Dept. of Information and Communication Engineering, Hanbat National University

**요약** 2018년부터 기후변화에 따른 폭염과 한파에 의해 사회기반시설에 점증적으로 많은 피해를 미치고 있다. 최근 4년간 기후변화에 따른 영향 중에 냉해에 대한 피해가 매년 증가하여 특정 지역에 국한되었던 피해가 이제는 전국에 걸쳐서 나타나고 있으며, 이를 최소화하기 위한 각계각층의 전문가들에 의한 노력이 집중되고 있다. 그러나 불연속적으로 측정되는 데이터들 속에서 지역 특색을 반영하고자 하는 기존 연구들에서는 갑작스럽게 발생하는 이상 저온에 대한 실시간 관측 연구는 쉽지 않은 상황이다. 본 연구에서는 냉해 발생에 영향을 미치는 기상 데이터를 기준으로 냉해 피해가 발생하였던 시점의 기상 패턴을 탐색한 후 이상저온 발생 시점을 확인할 수 있는 알고리즘 패턴 개발을 하였다. 본 연구의 결과물은 과거의 데이터에 의존하지 않고 실시간으로 발생하는 데이터에 의한 이상저온이 발생한 시점을 시간 시점에서 확인할 수 있다는 점에서 지역적 기상 편차를 의식하지 않고 확인이 가능하다는 점에서, 이상저온 발생 시점 데이터를 확보할 수 있다는 점에서 관련 분야에 많은 도움이 될 것으로 기대할 수 있다.

**키워드** : 머신러닝, 학습모델, 이상저온, 기후변화, 냉해, 빅데이터

**Abstract** Since 2018, due to climate change, heat waves and cold waves have caused gradual damage to social infrastructure. Since the damage caused by cold weather has increased every year due to climate change in recent 4 years, the damage that was limited to a specific area is now appearing all over the country, and a lot of efforts are being concentrated from experts in various fields to minimize this. However, it is not easy to study real-time observation of sudden abnormal low temperature in existing studies to reflect local characteristics in discontinuously measured data. In this study, based on the weather-related data that affects the occurrence of cold-weather damage, we developed an algorithm pattern that can identify the time when abnormal cold temperatures occurred after searching for weather patterns at the time of cold-weather damage. The results of this study are expected to be of great help to the related field in that it is possible to confirm the time when the abnormal low temperature occurs due to the data generated in real time without relying on the past data.

**Key Words** : Machine learning, Learning model, Abnormal low temperature, Climate change, Cold damage, Big data

## 1. 서론

전 세계적으로 큰 관심을 가지고 있는 기후변화는 화산 분화, 태양 활동의 변화 및 천문학적인 상대위차 변화 등과 같은 자연적인 원인에 의한 변화도 있겠지만, 이보다 더 큰 문제점은 산업의 발전에 따른 온실가스 배출 등의 인위적인 요소가 매우 큰 원인으로 자각하고 온실가스 배출량을 줄이기 위해 각 국가에서는 많은 노력을 기울이고 있다. 2018년부터 기후변화에 따른 폭염과 한파에 의해 생태계, 산업, 건강, 농·어촌 등의 사회기반시설에 점증적으로 많은 피해를 미치고 있다. 한반도의 전 지구적 현상에 따른 변화는 매우 빠른 속도로 진행되고 있으며, 지역적으로 편차가 있는 기후 상태의 변화에 따른 영향을 파악하는데 많은 관심이 집중되고 있다[1,2].

2020년 기상청에서는 기상 예측을 위한 알파웨더(Alpha Weather)를 개발하고 있었으나, 실제 기상 예측에는 사용되지 않았으며, 다만 미래의 날씨를 정교한 기후 모델을 통해 예측하는 자연법칙과 온도, 습도와 같은 결과 데이터를 이용하여 과거 기후 관측 데이터를 바탕으로 찾는 방식을 사용하고 있다. 결과 데이터 도출에 사용되는 과거 관측 데이터의 경우 불연속적으로 측정되는 오류를 범하고 있다. 실제로 기상 데이터의 경우에는 실시간 관측 데이터로 연속성을 가지고 있기 때문이다[3].

기존 연구의 경우에는 과거 데이터의 표준편차에 의한 표현으로 지역 특성을 반영하고 있기 때문에 갑작스럽게 발생하는 이상저온에는 대처하기 어려운 상황이다[4]. 즉, 매년 변동되는 실시간 기상 데이터에 의한 이상저온 발생 시점을 지역 특성과 무관하게 범용적으로 확인하기가 어렵다.

본 연구에서는 이상저온 발생에 영향을 미치는 기온, 습도, 풍속, 풍향, 기압 등의 기상 데이터 활용 및 실제 냉해 피해가 발생했던 일자의 기상 패턴을 탐색적 데이터 분석(EDA; Exploratory Data Analysis)을 활용하여 패턴에 맞는 알고리즘을 개발하여 이를 기준으로 이상저온이 발생한 시점(hour) 확인을 위한 분류 모델을 선정하고자 한다.

본 연구는 기후변화에 따른 다양한 변화들 중에 이상저온에 한정하여 연구를 진행하였다.

## 2. 관련연구

### 2.1 기후변화 정의

기후변화협약은 기후 변화에 관한 국제 연합 기본 협약(The United Nations Framework Convention on Climate Change. 약칭으로 유엔기후변화협약 또는 기후변화협약 또는 UNFCCC 또는 FCCC로 사용됨.)은 온실가스 등의 기체에 의해 발생하는 지구 온난화 현상의 지속적 저감을 위한 국제적 협약이다.

1992년 6월에 브라질 리우데자네이루에서 기후변화협약이 체결되었고, 각종 온실가스의 유출 및 이산화탄소를 제한하고 지구 온난화를 막는데 가장 큰 목적을 두고 있는 협약으로, 기후변화에 대한 정의를 전 지구 대기의 조성을 변화시키는 인간의 인위적인 활동이 직·간접적으로 원인이 되어 일어나며, 자연적인 기후 변동성에 추가하여 오랜 기간 동안 관측되어 일어나는 기후의 변화로 정의하고 있다[5].

### 2.2 기후변화 감시 문제점

기후 안보가 국제사회에 새롭게 이슈화 되고 있는 가운데 기후변화에 대한 현상 감시, 여건 변화에 맞는 표준화된 예측자료 및 예측 모델 개발을 통한 활용체계 구축에 대한 필요성이 증대되고 있다.

문제점으로는 기후변화 감시·예측 관련한 국내 현황 및 문제점, 기후변화에 대한 한국형 예측 모델 개발, 예측 자료의 생산, 현상 감시, 감시 예측정보의 활용체계 구축에 대한 감시 및 예측에 대한 문제점이 항상 존재하고 있기에 국내 기후변화에 대한 대응을 위하여 기상청은 좀 더 상세한 한반도 내 기후변화 시나리오에 대한 산출이 필요하다[6,7].

### 2.3 탐색적 데이터 분석 방법

탐색적 데이터 분석(EDA; Exploratory Data Analysis)은 데이터를 분석하고 조사하여 주요 특성을 파악하는 데에 주로 사용되는 기법으로, 데이터 시각화 방법을 사용하기도 한다. 탐색적 데이터 분석의 주요 절차는 문제를 해결하기 위해 수집 데이터를 파악하고 가공하여 이슈화 될 패턴을 찾고, 비정상적인 수치를 처리하며, 테스트 및 결과 확인 등의 절차로 분석하게 된다. 이처럼 탐색적 데이터 분석은 주로 모델의 정규화와 데이터 종류의 확인과 검증 작업을 포함하여 데이터 특성 간의 관계에 대한 직

관적인 이해를 제공하는 데에 사용된다. 또한 데이터 분석을 위해 사용할 기법이 적합한지 여부를 판단하는 데에도 매우 유용하게 사용된다. 탐색적 데이터 분석 기술은 현재 데이터 검색 프로세스에 널리 사용되고 있다[8-12].

### 2.4 상관관계 분석 방법

상관관계 분석(Correlation analysis)은 두 개 이상의 특성 간에 어떠한 관계가 있는지를 분석하는 기법으로 상관계수의 값을 이용하여 선형관계를 확인하게 된다. 두 개 이상의 특성들은 서로 상관없이 독립적인 관계이거나 서로 상관된 관계일 수도 있으며, 이때 두 개 이상의 특성 간에 발생한 관계에 대한 강도를 상관관계라고 한다[13,14].

### 2.5 분석유형에 따른 분석 모델

기계학습을 통한 분석은 대부분 분류와 회귀에 속한다. Table 1과 같이 분류는 범주형 데이터를 예측하는 것으로, '예' 또는 '아니오'와 같은 값의 형태로 분류하는 데에 사용되며, 회귀는 연속된 데이터를 예측하는 데에 사용된다[15].

**Table 1. Analysis model according to analysis type**

Analysis Method	Analysis Type	Analysis Model
Supervised Learning	Classification	KNN(K-nearest neighbors)
		Naive Bayes
		Decision Tree
		Logistic Regression
		Random Forest
		Support Vector Machine(SVM)
		Artificail Neural Network(ANN)
	Estimation	Linear Regression(Stepwise)
		Regularized Linear Regression
		Regression Tree
		Random Forest Regression
		Support Vector Regression
		Clustering
Unsupervised Learning	Clustering	Hierarchical Clustering
		K-means Clustering, K-medios
		Self-Organizing Map(SOM)
	Association	Market Masket Analysis(MBA)
		Sequence Analysis
		Collaborative Filtering
		Dimension Reduction
	Dimension Reduction	A Priori
		Principak Component Anaysis(PCA)
		Factor Analysis
		Multi-Dimensional Scalling(MDS)

### 2.6 데이터 ETL 프로세스

ETL(Extract, Transform, Load)은 데이터의 추출(Extract), 변환(Transform), 로드(Load)로 다양한 데이터를 가공하고 탐색하기 위해 사용되는 과정으로, 분석하고자 하는 비즈니스 주제에 따라 데이터를 변환하고 데이터를 저장하고 다시 읽어 들인다. 보통 데이터 변환에 사용되는 방법으로는 정렬, 필터링, 집계, 데이터 조인, 중복 제거, 데이터 정리 및 데이터에 대한 유효성 검사 등 다양한 전처리 및 가공 작업이 사용되며, 시간 절약을 위해 ETL 단계가 동시에 진행되기도 한다[16].

본 연구에서도 필터링, 정렬, 집계, 데이터 조인, 데이터 가공, 결측치 제거 및 변수 간의 유의성 검사 등의 작업이 진행되었다.

## 3. 연구방법

### 3.1 연구 범위

본 연구는 사전에 기상청 기상자료개방포털 담당자에게 이상저온 발생 시점 데이터 유무에 대한 문의 결과, “이상저온 발생 시점에 관한 데이터는 없으며, 이를 확인하기 위해서는 온라인 뉴스 매체 등의 냉해 발생 보도자료를 확인하는 방법이 좋을 것 같다”는 의견에 따라, 이상저온 발생 시점 확인을 위한 연구가 시작되었으며, 본 연구의 범위를 다음과 같이 정의하였다.

첫째, 일시적 이상저온에 따른 냉해 피해가 지속적으로 발생하고 있는 2018년부터 2021년까지의 뉴스 보도에서 피해 일자 및 지역을 선별하여 수집하였다.

둘째, 2018년부터 2021년까지 4년간 지속적으로 냉해 피해가 발생하였던 지역 중에 기상청의 기상자료개방포털에서 제공하는 종관기상관측 4년간의 데이터가 모두 존재하는 지역을 선정하여 기상 데이터를 수집하였다.

셋째, 실제로 냉해 피해가 발생하였던 일자에 대한 기온, 풍속, 습도, 풍향, 기압 등의 기상 데이터에서 어떠한 이상 패턴이 있는지를 탐색적 데이터 분석 방법을 통하여 탐색하였다.

넷째, 이상 패턴을 보이는 냉해 피해 발생 일자에 대한 알고리즘 및 조건을 개발 및 적용하여 이상저온 발생 시점 확인을 위한 조건으로 타당성이 있는지에 대한 정확도 검증을 진행하였다.

다섯째, 이상저온 발생 시점의 알고리즘 적용데이터를 기준으로 기온, 풍속, 습도, 풍향 및 기압 등의 기상 데이

터들을 요인변수(feature)로 사용하여 이상저온 시점 확인을 위한 분류 모델들의 학습을 통해 과대 및 과소적합이 일어나지 않는 일반화된 모델로 최종 선정하였다.

여섯째, 생성된 모델들의 테스트를 위하여 연구 지역 외에 외부 지역 한곳을 선정하여 4년간 전수 데이터를 모델에 적용하여 최종 확인하였다.

### 3.2 연구 지역 선정 및 데이터 수집

이상저온 발생 시점 연구를 위하여 Table 2와 같이 연도, 월, 일자 및 발생 지역과 Table 3, Table 4에서와 같이 2018년부터 2021년까지 4년간 냉해 피해가 발생하였던 지역에 대한 데이터를 수집하였다.

본 연구에서는 천안 지역의 데이터를 기준으로 이상저온이 발생했던 시점에 냉해 피해가 발생한 것을 밝힘으로써, 이상저온과 냉해 발생과의 연관성을 분석하고자 하였으며, 테스트 지역으로는 원거리 지역인 순천 지역을 기준으로 검증하였다.

Table 2. Synoptic meteorological data offer list

Division	Offer Details
Data Type	Year, Month, Day, Time, Minute
Offer Period	Since 1904
Offer Point	102 Points
Offer Element	Temperature, Precipitation, Wind, Air Pressure, Humidity, Solar Radiation, Sunshine, Snow, Clouds, Visibility, Ground Condition, Ground Temperature, Weather Phenomenon, Evaporation Amount, Development Number

Table 3. Status of areas affected by cold weather by year

Reference Year	Occurrence Zone
2018	Geumsan, Cheonan, Gongju, Nonsan, Buyeo, Seosan, Taean, Boeun, Cheongju, Muju, Imsil, Jangsu, Jinan, Geochang, Goryeong, Sangju, Yeongcheon, Ulsan, Naju, Suncheon, Anseong, Hwaseong, Gapyeong, Namyangju, Yeosu, Icheon
2019	Ulsan, Jinju, Hadong, Daegu, Cheonan, Chungju, Boeun, Okcheon, Jeongpyeong, Jincheon, Goesan, Anseong, Gimcheon, Naju, Yeongam, Suncheon, Sangju
2020	Suncheon, Cheonan, Yeongdong, Okcheon, Andong, Bonghwa, Hadong, Anseong, Icheon, Naju, Jangsu, Jinan, Muju, Jeonju, Ulsan, Geochang, Hamyang, Yeongcheon, Pohang, Bonghwa, Cheongsong, Sangju, Jinju, Boeun
2021	Yeongam, Naju, Bonghwa, Yeongju, Anseong, Goesan, Eumseong, Cheonan, Yeongam, Wanju, Suncheon, Jangheung, Hwasun, Yeongam, Haenam, Gangjin, Boseong, Goheung, Jangseong, Gurye, Gokseong, Damyang

Table 4. Status of areas affected by cold weather by the number of incidents over four years

Number of Occurrences	Occurrence Zone
1	Gapyeong, Gangjin, Goryeong, Goheung, Gokseong, Gongju, Gurye, Geumsan, Gimcheon, Namyangju, Nonsan, Damyang, Daegu, Boseong, Buyeo, Seosan, Seongju, Andong, Yeosu, Yeongdong, Yeongju, Wanju, Voice, Imsil, Jangheung, Jeonju, Jeungpyeong, Jincheon, Cheongsong, Cheongju, Chungju, Taean, Pohang, Hamyang, Haenam, Hwaseong, Hwasun
2	Geochang, Goesan, Muju, Yeongcheon, Okcheon, Icheon, Jangsu, Jinan, Jinju, Hadong
3	Boeun, Bonghwa, Sangju, Ulsan
4	Cheonan, Suncheon, Naju, Anseong, Yeongam

### 3.3 환경 데이터 간의 상관관계 분석

환경 변수들 간에 상관관계 분석 결과 Table 5와 같이 기온과 기압과의 상관계수가 0.77로 가장 높게 분석되었으며, 습도와 풍속, 습도와 풍향, 풍속과 풍향과의 상관계수는 모두 0.5 이상으로 상관관계가 있다는 것으로 분석되었다.

Table 5. Synoptic meteorological data offer list

	Temperatures	Wind Speed	Wind Direction	Humidity	Atmospheric Pressure
Temperatures	1.000	0.236	0.088	-0.030	-0.773
Wind Speed	0.236	1.000	0.660	-0.558	-0.158
Wind Direction	0.088	0.660	1.000	-0.510	-0.041
Humidity	-0.030	-0.558	-0.510	1.000	-0.181
Atmospheric Pressure	-0.773	-0.158	-0.041	-0.181	1.000

### 3.4 냉해 발생 요인 분석

냉해 발생 기준 데이터를 포함하여 기온, 습도, 풍속, 풍향, 기압에 대한 요인들에 대하여 선형회귀 분석을 통해 확인할 수 있으며, 이에 대하여 선형회귀 분석 중 여러 요소를 포함하는 다중회귀분석을 통해 냉해 발생 예측 요소로 사용하는데 유의미한지에 대하여 아래 그림 4.2.1, 그림 4.2.2, 그림 4.2.3, 그림 4.2.4와 같이 다중회귀분석 방법으로 확인하였다.

다중회귀분석 방법의 경우에는 전진선택방법(Forward Selection Method), 후진소거방법(Backward Elimination Method), 단계적선택방법(Stepwise Selection Method) 등 3가지 방법으로 수행하였으며, 3가지 방법 중 가장 높은 수정결정계수(Adjusted R-squared)를 채택하여 요소들의 유의미 여부 및 예측 가능성을 확인해 보았다.

냉해 발생 예측 데이터를 기준으로 기온, 습도, 풍속,

풍향, 기압, 냉해예측기준의 모든 요소들이  $p < 0.05$ 로 유의미한 관계가 있다는 것을 확인하였다.

### 3.4 탐색적 데이터 시각화 패턴 개발

4년간의 이상저온에 따른 냉해 발생 일자에 대한 탐색적 데이터 시각화 분석을 통하여 환경 변수들 간에 패턴을 확인할 수 있었으며, 냉해 발생 일자에 대한 특정 시점(시간)에 기온은 낮고, 습도는 높으며, 풍향은 0, 풍속은 0.4보다 작으며, 기압은 상대적으로 음의 상관관계 패턴으로 분석되었다. Fig. 1 및 Fig. 2는 4년간 데이터 시각화 분석의 패턴이 모두 동일하기에 2018년 1개년도에 대한 냉해 발생 일자에 대해서만 대표적으로 나타내었다.

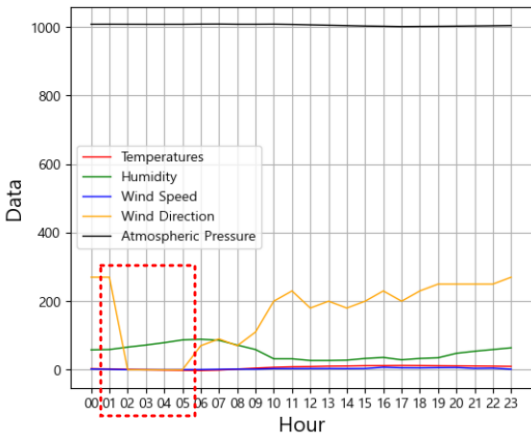


Fig. 1. Analysis of weather data patterns at the time of occurrence of abnormal low temperature by time zone

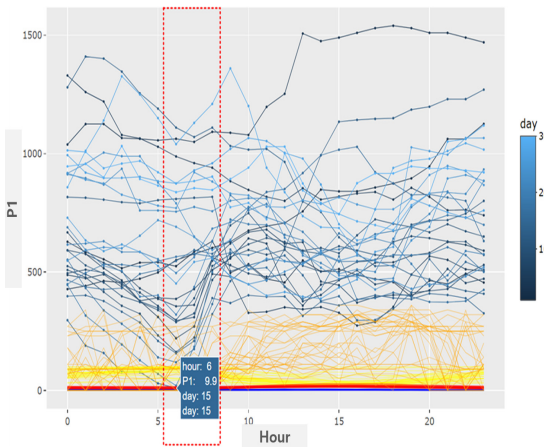


Fig. 2. Changes based on the time of occurrence of cold damage by time zone

- 이상저온 발생 알고리즘
  - 이상저온 기준(P1) = 기온(t)×습도(h)
  - 이상저온 기준(P2) = 기온(t)×습도(h))/100
- 이상저온 발생 패턴
  - ① 이상저온 기준(P1) < |기온(t)|
  - ② 이상저온 기준(P2) < 습도(h)
  - ③ 풍속(ws) <= 0.4
  - ④ 풍향(wd) == 0

위 ①, ②, ③, ④의 조건을 모두 만족할 경우 이상저온에 의한 냉해가 발생한다.

### 3.5 이상저온 발생 시점 확인 학습모델 선정

연구 지역인 천안 지역의 기상 데이터에서 이상저온 발생 패턴을 보이고 있는 것으로 분석되었다. 이에 따라 이상저온 발생 시점을 확인하기 위하여, 이상저온 발생 및 미발생에 대한 기계학습 분류 분석을 하였다.

Table 6과 같이 정확도, 정밀도, F1-스코어 및 재현율의 값이 높으며, 최대적합 및 최소적합이 일반화된 모델로 랜덤포레스트(RandomForest) 모델을 선정하였다.

- 랜덤포레스트 분류 예측 모델을 통한 전수 데이터 검증
  - RandomForestClassifier(n\_estimators=n,max\_depth=depth)

Table 6. Train and verification model accuracy list

Model	Train Accuracy	Verification Accuracy	Recall
Random Forest	0.999	0.998	0.998
ExtraTrees	0.986	0.985	0.830
Gradient Boosting	1.000	1.000	1.000
Histogram-base Gradient Boosting	1.000	0.999	0.996
XGBoost	1.000	1.000	1.000
LightBGM	1.000	0.999	0.998

위 결과는 분류 분석 기계학습 모델인 랜덤포레스트(RandomForest), 엑스트라트리(ExtraTrees), 그래디언트부스팅(Gradient Boosting), 히스토그램 기반 부스팅(Histogram-base Gradient Boosting), XGBoost, LightBGM 등 6개의 훈련 모델에 기상 데이터 정규화 및 트리 개수, 최대 깊이, 분리 노드의 최소 자료수, 잎사귀

노드의 최소 자료수에 대한 하이퍼파라미터 튜닝 결과를 각 모델에 적용하여 분류 분석을 하였다.

### 4. 연구 결과

#### 4.1 이상저온 발생 시점 확인 모델

이상저온 발생 시점을 확인하기 위한 기계학습 모델 중에 최종 선정된 램덤포레스트 모델은 아래와 같다.

- RandomForestClassifier(max\_depth=depth, min\_samples\_leaf=leaf, min\_samples\_split=split, n\_estimators=n)

#### 4.2 선정 모델을 활용한 타 지역 이상저온 발생 시점 확인

선정한 이상저온 발생 시점 확인 모델을 활용하여 타 지역인 순천 지역의 기상 데이터를 적용한 결과로는 Table 7 및 Fig. 3의 Confusion Matrix 결과와 같이 기존의 과거 데이터 및 지역적 특성을 고려하지 않더라도 매우 높은 성능을 보이는 모델로 확인되었다.

Table 7. Finally test model accuracy

Test Accuracy	Precision	Recall	F1-Score
0.999	0.980	0.997	0.989

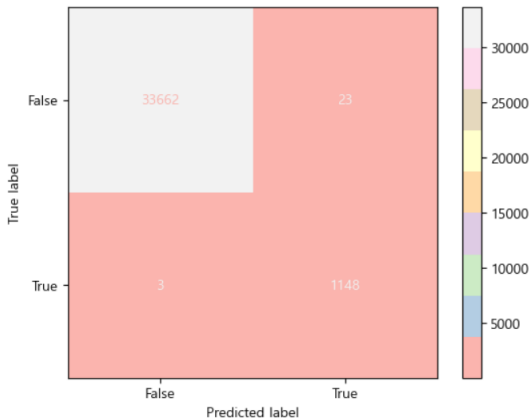


Fig. 3. Confusion matrix display

### 5. 결론

본 연구의 결과는 기후변화에 대응하는 실시간 기상 데이터를 적용하여 이상저온 발생 시점을 확인할 수 있는

모델을 선정하였다는 점에서 지역별 기상 데이터의 차별화 등 기상변화와 이상저온의 변화에 큰 영향을 받지 않으며, 기존의 평균값에 의존하였던 분석에서 더 나아가 실시간 데이터를 기준으로 분석했을 때 의미가 더 큰 모델이라는 점에서 의미가 있는 연구이다.

본 연구 모델로 기상청에서 제공되는 3일 또는 5일 전 기상예보 데이터를 적용한다면, 이상저온 발생 일자 및 시점을 사전에 확인할 수 있으며, 사전 확인을 통해 산림, 농·어촌 지역에 이상저온에 의한 피해를 최소화하는데 기여할 수 있다.

또한, 기계학습에 기반한 연구 모델을 딥러닝(Deep Learning) 기반으로 모델을 연구 한다면 기계 학습으로 찾지 못했던 또 다른 특성 및 패턴을 찾아낼 수 있을 것으로 판단된다.

### REFERENCES

- [1] Ministry of Environment, "Korea Climate Change Assessment Report 2020," 2020.
- [2] NATIONAL INSTITUTE OF METEOROLOGICAL SCIENCES, "Response to IPCC 6th Assessment Report, Global Climate Change Prospect Report," 2019.
- [3] IT Chosun. In the age of AI and big data, why is the Korea Meteorological Administration wrong?. Retrieved from [http://it.chosun.com/site/data/html\\_dir/2020/08/05/2020080501239.html](http://it.chosun.com/site/data/html_dir/2020/08/05/2020080501239.html)
- [4] D. J. Kim, "Categorizing Likelihood of Agrometeorological Hazards into Risk Indices," *Korean Society of Agricultural and Forest Meteorology*, 2014, pp. 61-73.
- [5] Korea Environmental Institute. Korea Adaptation Center for Climate Change, Korea Adaptation Portal for Climate Change. Retrieved from [https://kacc.kei.re.kr/portal/climateChange/climatechange\\_list.do](https://kacc.kei.re.kr/portal/climateChange/climatechange_list.do).
- [6] Korea Environmental Institute. Korea Adaptation Center for Climate Change, Korea Adaptation Portal for Climate Change. Retrieved from [https://kacc.kei.re.kr/portal/climateChange/changeeffect/changeeffect\\_view.do?num=8](https://kacc.kei.re.kr/portal/climateChange/changeeffect/changeeffect_view.do?num=8).
- [7] Korea Environmental Institute. Korea Adaptation Center for Climate Change, Korea Adaptation Portal for Climate Change. Retrieved from

[https://kaccc.kei.re.kr/portal/climateChange/changepheno/changepheno\\_view.do?num=3](https://kaccc.kei.re.kr/portal/climateChange/changepheno/changepheno_view.do?num=3).

- [8] IBM, "Exploratory Data Analysis-Definition and Tools," 2021.
- [9] Wiki Pedia Exploratory Data Analysis. Retrieved from [https://ko.wikipedia.org/wiki/Exploratory Data Analysis](https://ko.wikipedia.org/wiki/Exploratory_Data_Analysis).
- [10] M. H. Heo & J. H. Jeong (1990). Comparative Review of Statistical Package Programs for Exploratory Data Analysis Functions: *The Korean Statistical Society*, 3(2), 17-25.
- [11] Y. H. Shin (2021). Data Science Hard Carry - For the Age of Artificial Intelligence: Life & Power Press.
- [12] Y. H. Shin (2021). Prediction Model for Pre-Sale Price of Multi-Unit Housing Development Project Based on Machine Learning: Kyung Hee University Doctoral Dissertation.
- [13] Wiki Pedia. Correlation analysis. Retrieved from [https://ko.wikipedia.org/wiki/Correlation analysis](https://ko.wikipedia.org/wiki/Correlation_analysis).
- [14] Bioinformaticsandme. Correlation analysis. Retrieved from <https://bioinformaticsandme.tistory.com/58>.
- [15] Y. J. Bae, H. C. Kang, K. P. Yeon, E. M. Kwon & S. T. Han (2015). Comparison and Case Analysis of Cluster Analysis Algorithms for Mixed Quantitative and Qualitative Data: *The Korea Data Analysis Society*, 17(6), 2991-3002.
- [16] Microsoft Docs, ETL. Retrieved from <https://docs.microsoft.com/ko-kr/azure/architecture/data-guide/relational-data/etl>.

이정원(Jeongwon Lee)

[정회원]



- 2009년 2월 : 공주대학교 교육대학원 컴퓨터교육과(교육학석사)
- 2014년 8월 : 공주대학교 대학원 컴퓨터교육과(박사수료)
- 2022년 8월 : 한밭대학교 정보통신전문대학원 정보통신공학과(공학박사)
- 2016년 10월~현재 : 빅데이터·정보보호융합기술협회 이사
- 관심분야 : 빅데이터, 인공지능, 데이터베이스, 웹&응용프로그래밍, 소프트웨어교육
- E-Mail : mentor1023@daum.net

이충호(Choong Ho Lee)

[정회원]



- 1985년 2월 : 연세대학교 전자공학과(공학사)
- 1987년 2월 : 연세대학교 대학원 전자공학과(공학석사)
- 1998년 3월 : 도호쿠대학 대학원 정보과학연구과(공학박사)
- 1987년 2월~2000년 2월 : KT 멀티미디어연구소 전임연구원
- 2000년 2월~현재 : 한밭대학교 정보통신공학과 교수
- 관심분야 : 영상처리, 컴퓨터비전, 기계학습, 빅데이터, 소프트웨어교육
- E-Mail : chlee@hanbat.ac.kr