



## 머신러닝을 활용한 산불발생예측모형 개발과 과제

# Machine Learning for Big Data Analytics in Development of Wildfire Prediction Models

이찬중\* · 임무영\*\* · 이요한\*\*\*

Lee, Chanjung\*, Lim, Mooyoung\*\*, and Lee, Yohan\*\*\*

### Abstract

This study aims to develop a model that predicts domestic forest fire occurrences during fire outbreaks using machine learning techniques. For the modeling methods, logistic regression analysis and ensemble techniques, such as gradient boost and random forest, were used while the oversampling technique was utilized to address the imbalance problem of the forest fire data. The model developed in this study predicted 239 out of 333 forest fire occurrences during the nationwide forest fire period in 2020 with a prediction accuracy of approximately 71.8%. Forest fires that occur during such periods are highly influenced by different factors affecting the climate, such as temperature, humidity, and precipitation. In Gangwon-do, in addition to these factors, a high correlation between farmland density and stem volume per hectare has also been associated with increased forest fire occurrences. The significance of this study lies in the fact that it presents a customized wildfire occurrence prediction model that can be used in the administrative parts, which serve as the basic centers for wildfire prevention, of provinces and cities across the country.

**Key words :** Machine Learning, Imbalanced Data, Forest Fire Occurrence, Prediction, Ensemble Model

### 요 지

본 연구의 목적은 머신러닝 기법을 활용하여 산불기간 동안 국내 산불발생을 예측하는 모형을 개발하는 것이다. 모델링 방법으로는 로지스틱 회귀분석과 Gradientboost, Randomforest 등의 앙상블 기법을 활용하였으며, 산불데이터의 불균형 문제를 해결하기 위하여 오버샘플링 기법을 활용하였다. 본 연구에서 개발한 모형을 활용하여 2020년 전국 산불기간 산불발생을 예측한 결과 전체 333건의 산불 중 239건의 산불을 예측했으며 예측정확도는 약 71.8%로 나타났다. 전국적으로 산불기간 동안 발생한 산불은 기온, 습도, 강수량 등 기후 요인에 큰 영향을 받는 것으로 보였다. 반면 강원도에서는 기후요인 이외에도 농지 밀도와 ha당 재적이 산불발생과 높은 상관관계가 있는 것으로 나타났다. 본 연구는 두 가지 측면에서 의의를 가진다. 첫째, 산불발생을 시·군·구 단위의 지역별로 나타내면 산불기간 동안 산불이 발생하는 지역보다 발생하지 않는 지역이 많아 데이터 값의 불균형이 나타난다. 본 연구에서는 오버샘플링 기법을 통해 데이터의 불균형을 해소하는 방안을 제시하였다. 둘째, 국내 산불정책은 행정구역으로 구분하여 예방 및 진화활동을 하고 있는데, 이를 위해 행정구역 단위의 산불발생 예측 및 대책 마련이 필요하다. 본 연구는 행정구역 단위에서 적용 가능한 산불발생예측모형을 제시하였다는데 중요한 의의가 있다.

**핵심용어 :** 머신러닝, 불균형데이터, 산불발생, 예측모형, 앙상블 모델

\*정회원, 서울대학교 농업생명과학대학 농림생물자원학부 박사과정(E-mail: gtano@snu.ac.kr)

Member, Ph.D. Student, Dept. of Agriculture, Forestry and Bioresources, Seoul National University

\*\*경북대학교 농업생명과학 산림과학·조경학부 입학과 석사과정

Master Student, School of Forest Sciences and Landscape, Kyungpook National University

\*\*\*교신저자, 정회원, 서울대학교 농업생명과학대학 농림생물자원학부 부교수(Tel: +82-2-880-4754, Fax: +82-2-873-3560, E-mail: johnlee@snu.ac.kr)

Corresponding Author, Member, Associate Professor, Dept. of Agriculture, Forestry and Bioresources, Seoul National University

## 1. 서론

기후변화에 따라 고온 건조한 날씨가 지속되면서 전 세계적으로 산불발생이 증가할 뿐 아니라 피해의 규모도 커지고 있다(KFS, 2009; Pechony and Shindell, 2010). 2022년 국내에서만 740건의 산불이 발생하였고, 총 피해면적은 24,782 ha에 달한다. 특히, 2022년 3월에 발생한 울진-삼척 산불의 피해면적은 20,923 ha에 달하였고, 총 213시간 동안 산불이 이어지면서 역대 최장 산불로 기록되었다(Hankyoreh, 2022). 울진-삼척 산불의 직접피해액은 1,717억원으로 최종 집계되었고, 3,027억원의 복구비가 발생하였다. 산림이 가지는 다양한 공익적 기능을 고려하면 2022년 산불발생으로 인한 전체 피해액은 약 1.3조원에 이를 것으로 추산한다(Lee et al., 2022).

우리나라 산불은 기온 및 강수량 등과 같은 기상적 요인에 의해 영향을 많이 받아 계절별 발생 빈도의 차이가 뚜렷한 편이다. 2022년 산불통계 자료에 의하면, 최근 10년간(13~22) 연평균 약 535건의 산불이 발생하였고 이 중 약 65%가 봄철에 집중적으로 발생하였다. 월별 산불발생 빈도는 3월(129건/년), 4월(119건/년), 2월(70건/년) 순으로 나타났다(KFS, 2023). 우리나라의 산불발생은 기상적 요인 외에도 인위적 요인에 의해서도 영향을 받는데, 이는 산불이 자연적으로 발화되기보다는 산림에서 사람들의 활동으로 주로 발생하기 때문이다. 인위적 요인 중에서도 입산자 실화, 소각산불에 의해 발생한 산불이 최근 10년간 발생한 산불 전체의 46%를 차지한다(KFS, 2018, 2023).

과거 국내에서 산불발생에 영향을 주는 인자와 산불발생 간의 상관관계 혹은 인과성을 분석하는 연구가 활발하게 이루어져 왔으며(Lee and Lee, 2006; Kwak et al., 2010), 산불의 초기 대응과 연계할 수 있는 산불발생예측에 관한 모형도 다양한 연구자들에 의해 제시되었다. Lee et al. (2004)은 국내 지역별 산불발생에 영향을 주는 기상변수를 분석하여 실효습도, 일최고온도, 평균풍속 등 기상변수의 변동이 산불발생확률을 증가시킨다고 제안하였다. Kwak et al. (2010)은 산불발생의 공간분포가 인간 활동과 연관을 가진다는 가정 하에 산불발생 원인별 공간분포 형태를 파악하여 군집화 경향을 밝혀냈다. Lee et al. (2011)은 국내 과거 산불 자료를 기반으로 계절별 산불발생일, 일일 산불발생 건수, 산불발생 시간 등에 적합한 확률분포를 추정하여 우리나라 계절별 특성을 고려한 산불발생 확률 시뮬레이션 모델을 개발하였다.

최근 산불발생예측모형에 관한 연구는 주로 두가지 유형으로 구분이 가능하다. 첫 번째 유형은 산불발생의 임의적 특성에 착안하여 확률적인 산불 모형을 가정하여 시뮬레이션하는 방식이다(Fried et al., 2006; Lee et al., 2013). 다른 유형은 모형을 처음부터 가정하지 않고 과거 데이터를 가지고 모형을 학습시켜 산불발생을 예측하는 머신러닝 기법

등의 다양한 모델링 방식이다(Chae et al., 2018; Jain et al., 2020). 해외 연구를 살펴보면, Cunningham and Martell (1973)은 북서 온타리오에서 발생한 인위적 요인에 의한 일별 산불발생을 그 지역의 연료의 수분함량에 기반한 일별 산불발생건수에 대한 포아송 확률 프로세스를 이용하여 모델을 개발하였다. 이후 이러한 확률모형에 기반한 인위적 산불의 발생건수를 예측하는 연구들이 다수 이루어졌다(Martell et al., 1987; Martell et al., 1989; Todd and Kourtz, 1991; Pew and Larsen, 2001). 이들 연구에서는 연료의 수분함량과 인위적 산불발생 사이의 강한 상관관계를 공통적으로 발견하였다.

2000년대 이후에는 IT 기술의 발달로 가용한 데이터의 양이 혁신적으로 늘어났고, 컴퓨터의 계산 속도도 크게 향상되면서 빅데이터 기반의 머신러닝을 이용한 산불발생예측 모형에 대한 연구가 활발하게 이뤄지고 있다(Jain et al., 2020). Alonso-Betanzos et al. (2003)는 스페인의 실시간 산불관리시스템의 한 기능으로서 기온, 습도, 강우, 과거산불발생패턴의 정보를 포함하는 일일 산불발생 위험 지표를 예측하는 인공 신경망(Artificial Neural Networks, ANN) 모형을 개발하였다. Vasilakos et al. (2009)는 그리스 레스보스 섬지역의 산불발생을 예측하기 위해 기존의 ANN 모형을 확장하여, 산불 기후 지표(fire weather index), 산불잠재요인(fire hazard index), 산불발생가능성(fire risk index)을 포함하는 산불발화지표(fire ignition index)를 개발하였다. Sakr et al. (2010)은 특정 일에 발생하는 산불의 발생 건수를 예측하기 위해 SVM(Support Vector Machine)을 이용하여 일일 산불위험지수를 개발하였다. Sakr et al. (2010)은 ANN 모형이 일반적으로 SVM 모형보다 성능이 높게 나타나지만, SVM 모형이 산불의 발생 여부를 판단하는 이진법(binary) 체계에 더 적합하다고 제안하였다. 하지만 여전히 ANN 모형이 정보의 이용과 결과 산출에 용이한 것으로 보고되고 있으며, 특히 RNN 모형이 가장 예측 성능이 높은 것으로 제안되었다(Dutta et al., 2013).

최신 산불분야에 적용된 머신러닝 모형 개발에서는 앙상블 기법을 활용한 모델이 가장 포괄적으로 산불예측에 활용되고 있다. Stojanova et al. (2012)는 다양한 변수를 활용하여 산불을 예측하는 KNN, DT, LR, SVM, BN 등의 머신러닝 모델을 비교하여 평가하였다. 앙상블 모델인 DT (decision trees)와 RF (Random Forest)가 가장 높은 성능을 나타냈다. Vecin-Arias et al. (2016)은 RF 모델이 이베리안 반도의 낙뢰로 인한 산불발생을 예측하는데 있어서 선형회귀모형보다 뛰어난 성능을 보인다고 제안하였다. Van Beusekom et al. (2018)은 푸에르토 리코에서 산불발생을 예측하기 위해서 RF 모형이 적합하며, 강우량이 가장 중요한 예측 인자라고 보고하였다. 산불발생의 위험을 예측하기 위해 머신러닝을 이용한 모델링 기법은 국내외에서 지속적으로 진화하고 있다(Jain et al., 2020).

최근 기후변화 영향 등으로 산불위험 요인은 다변화되고

있는 한편, 발생 예측은 어려워지고 있으며(KFS, 2021), 산불발생을 예측하는 모형에서 인위적인 요인의 중요성이 커지고 있는 실정이다(Loepfe et al., 2011; Rodrigues and De la Riva, 2014). 따라서 산불발생을 예측하는 데 있어 기상적 요인은 물론 인위적인 요인과 관계된 다양한 변수들을 고려해야 신뢰할 수 있는 예측모형을 기대할 수 있다.

본 연구의 목적은 머신러닝 기법을 활용하여 산불기간 동안 국내 산불발생을 예측하는 모형을 개발하는 것이다. 모델링 방법으로는 로지스틱 회귀분석과 SVM, 그리고 Gradientboost, Randomforest 등의 앙상블 기법을 활용하였으며, 산불데이터의 불균형 문제를 해결하기 위하여 오버샘플링 기법을 활용하였다. 본 연구에서는 최신 산불 관련 빅데이터를 활용한 전국단위 산불발생예측모형을 개발하고, 산불이 계절별, 지역별로 편중하여 발생하는 특성을 고려하여 분석자료가 가지는 불균형성을 오버샘플링 기법을 활용하여 해소하고자 하였다. 본 연구의 결과는 향후 산불정책입안자들이 정밀하게 산불발생을 예측하여 효율적으로 산불 예방 및 진화 계획을 수립하는데 기여할 것으로 기대한다.

## 2. 산불동향

산림청에서는 산불발생 지역을 서울부터 제주까지 시·도급 17개 행정구역을 기준으로 구분하고 있으며 각각에 대한 산불발생 건수와 피해면적을 매년 집계하고 있다. 최근

10년간('13~'22) 지역별 산불발생 동향을 살펴보면, 연평균 약 535건의 산불이 발생하였으며, 그로 인한 피해면적은 연평균 3,559 ha였다. 2022년은 이례적으로 가장 많은 산불 발생 건수(740건)와 피해면적(24,795 ha)이 보고된 해로 기록되었다(Table 1). 산불발생 동향을 지역별로 살펴보면 최근 10년간 산불발생 건수 기준으로 경기(116건/년), 경북(89건/년), 강원(75건/년) 순으로 많은 산불이 발생하였으며 세 지역의 합계는 전체 산불발생 건수의 52%를 차지하였다. 피해면적 기준으로는 경북(2,064 ha/년), 강원(1,077 ha/년), 경남(170 ha/년) 순이었으며, 세 지역의 합계는 전체 산불 피해면적의 93%를 차지하였다. 최근 5년간('18~'22) 산불발생동향과 관련하여 특기할 사항으로는, 2020년 경기지역에서 가장 많은 산불이 발생하였다는 것과 피해면적이 1,000 ha가 넘는 산불이 강원(2019년, 2022년), 경북(2020년, 2022년), 경남 지역(2022년) 등에서 꾸준히 발생해 오고 있다는 점이다(KFS, 2023).

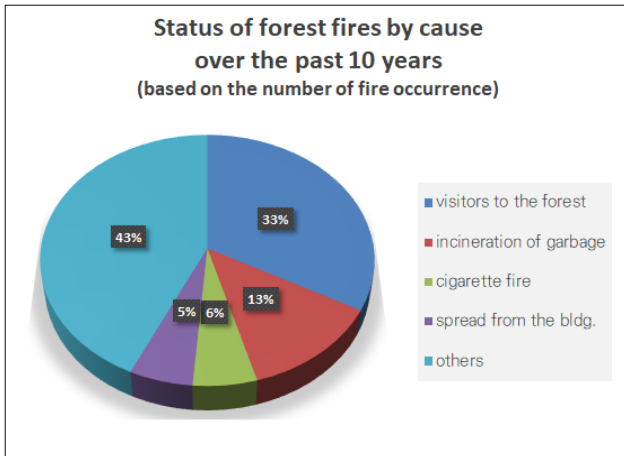
지난 10년간('13~'22) 원인별 산불발생 현황을 발생 건수 기준으로 살펴보면, 기타원인을 제외하면 입산자실화에 의한 발생이 176건/년(33%)으로 가장 큰 비율을 차지했고 쓰레기소각에 의한 발생이 68건/년(13%), 그리고 담뱃불실화와 주택화재비화에 의한 발생이 각각 30건/년(6%), 30건/년(5%)를 차지하였다(Fig. 1). 이 중에서 입산자 및 성묘객에 의한 실화가 전체 산불발생 원인에서 차지하는 비중은 최근 증가하고 있는 추세다. 한편 피해면적을 기준으로 산불발생 현황을 살펴보면, 기타원인(77%)의 비중이 커진 것을 확인

**Table 1.** Trends in Forest Fires by Province Over the Past 10 Years (2013~2022)

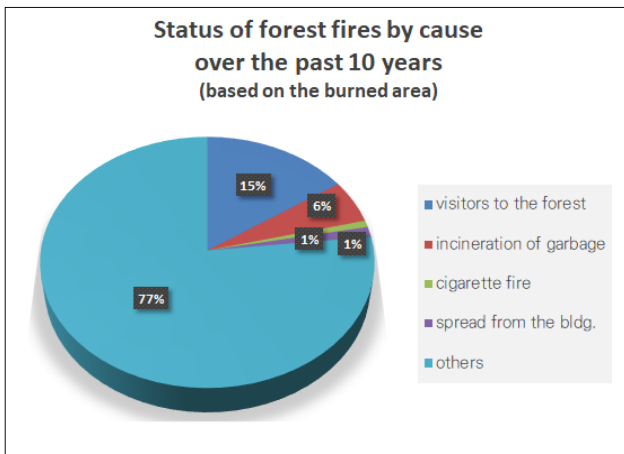
Province	2013		2014		2015		2016		2017		2018		2019		2020		2021		2022		10-yr avg.	
	occur.	area.	occur.	area.	occur.	area.	occur.	area.	occur.	area.	occur.	area.	occur.	area.	occur.	area.	occur.	area.	occur.	area.	occur.	area.
Seoul	4	0	19	1	21	2	13	1	16	3	16	0	6	1	4	1	4	0	9	3	11	1
Busan	15	1	11	3	11	1	8	1	13	1	15	66	22	73	16	1	6	1	21	30	14	18
Daegu	2	0	13	1	11	1	6	0	13	1	5	1	6	3	16	1	8	1	24	137	10	14
Incheon	4	0	22	3	44	6	25	12	21	5	19	1	24	3	14	12	8	1	8	1	19	4
Gwangju	7	1	7	1	5	0	1	0	2	0	3	1	2	0	2	0	1	0	4	0	3	0
Daejeon	6	1	9	2	12	5	3	5	18	4	5	0	9	2	3	0	4	1	3	0	7	2
Ulsan	14	320	13	2	7	2	4	0	11	19	12	5	13	2	18	531	11	3	30	20	13	90
Sejong	4	3	4	5	1	0	1	0	2	0	6	1	7	1	4	0	0	0	7	1	4	1
Gyeonggi	6	2	48	5	150	94	113	52	161	55	69	16	172	41	213	53	74	13	154	70	116	40
Gangwon	36	10	73	20	125	237	91	51	103	1,196	47	674	78	3,001	74	220	49	90	78	5,273	75	1,077
Chungbuk	14	4	18	2	31	7	19	78	46	8	33	1	28	4	13	4	8	76	24	16	23	20
Chungnam	17	17	23	6	23	4	15	2	52	16	29	3	39	4	35	4	24	41	51	186	31	28
Jeonbuk	32	13	29	6	21	3	13	1	24	3	29	4	23	3	23	3	23	16	51	37	27	9
Jeonnam	53	35	73	26	51	13	23	2	33	4	39	6	37	13	36	8	32	18	56	63	43	19
Gyeongbuk	45	118	98	48	87	35	46	172	115	154	100	106	121	95	106	2,053	60	449	114	17,409	89	2,064
Gyeongnam	35	26	32	6	23	9	10	2	62	10	69	8	66	8	41	26	37	56	105	1,550	48	170
Jeju	2	2	0	0	0	0	0	0	0	0	0	0	0	0	2	2	0	0	1	0	2	1
Total	296	552	492	137	623	418	391	378	692	1,480	496	894	653	3,255	620	2,920	349	766	740	24,795	535	3,559

\*'occur.' means the number of fire occurrence and 'area.' means the burned area in hectare

\*\*Source : Forest fire statistical yearbook (KFS, 2021), and Status of forest fire occurrence (KFS, 2023)



**Fig. 1.** Status of Forest Fires by Cause Over the Past 10 Years (Based on the Number of Fire Occurrence)



**Fig. 2.** Status of Forest Fires by Cause Over the Past 10 Years (Based on the Burned Area)

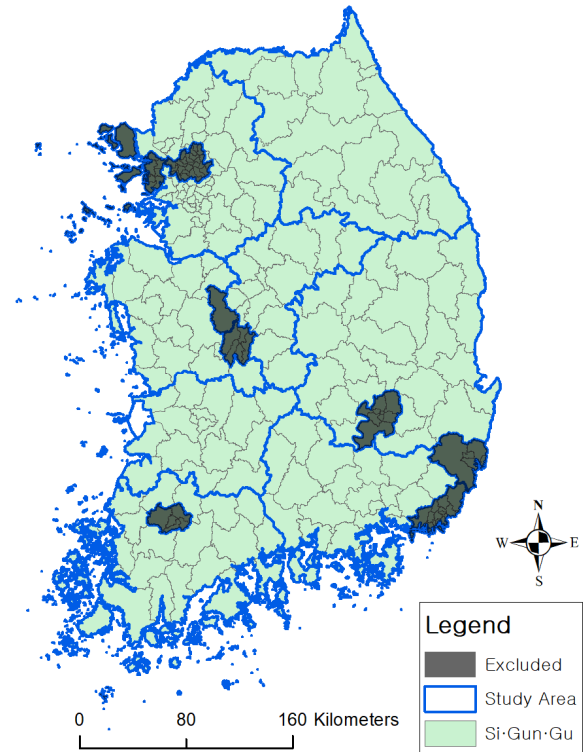
할 수 있으며, 그 뒤로 입산자실화(15%), 쓰레기소각(6%), 담뱃불실화(1%), 주택화재비화(1%) 순으로 산불발생 원인으로 파악되었다(Fig. 2). 이 통계는 2022년에 발생한 강원 지역과 경북 및 경남 지역의 초대형 산불을 포함하고 있어, 2022년 이전의 국내 산불발생 추이와는 차이가 있을 수 있다.

### 3. 연구방법

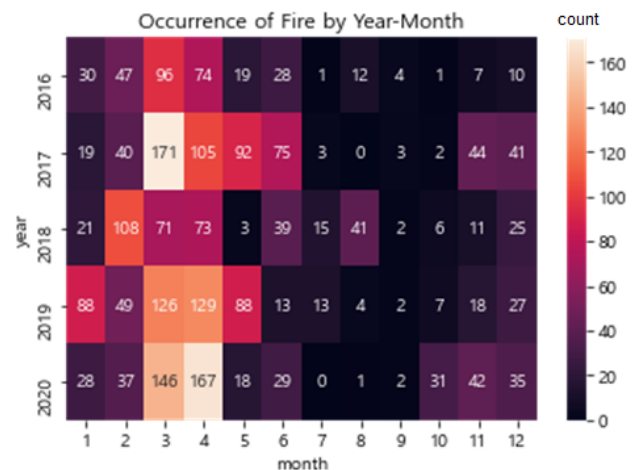
#### 3.1 연구 대상 및 범위

본 연구에서는 특별시·특별자치도·광역시를 제외한 전국의 도급 8개 행정구역(경기도, 강원도, 충청북도, 충청남도, 전라북도, 전라남도, 경상북도, 경상남도)을 연구대상지로 하였으며, 각 도에 존재하는 모든 시·군·구를 분석 대상으로 하였다(Fig. 3).

Fig. 4는 연구대상지 내에서 2016년부터 2020년까지 연도



**Fig. 3.** Study Area and the Level of Analysis



**Fig. 4.** Forest Fire Occurrence by Year and Month Restricted in the Study Area

별, 월별 산불이 발생한 빈도를 히트맵 형태로 보여주고 있다. 해당 기간 내 연도별, 월별 산불발생빈도가 가장 높은 경우는 2017년 3월(171회), 2020년 4월(167회), 2020년 3월(146회) 순이었으며, 대체로 2~4월에 집중되어 있었다. 따라서 본 연구의 시간적 범위는 2016년부터 2020년까지로 하였으며, 그중에서도 산불이 가장 많이 발생하였던 2~4월로 한정하였다.

### 3.2 연구 흐름

본 연구는 아래 순서(Fig. 5)에 따라 진행되었다. 먼저는 산불발생예측모형에 필요한 변수들을 선정하고 해당 변수들을 관련 기관으로부터 수집한 뒤 정제하고, 모델학습에 필요한 구조로 변환하여 데이터프레임 형태로 자료를 구축하였다. 다음으로는 모델학습과 모델평가를 위한 변수들을 분리하고 학습데이터를 임의 추출하여 학습에 필요한 최종 데이터를 선정하였다. 모델학습을 위하여 1개의 로지스틱 회귀 모형(Logistic Regression)과 1개의 서포트벡터머신(Support Vector Machine) 분류모형과 3개의 앙상블 모형을 사용하였다. 이때 8개의 행정구역에 대하여 6개 모형을 각각 생성하였으며, 모델학습에 필요한 학습데이터와 검증데이터의 분할은 8:2로 설정하였다. 모델의 평가를 위하여 2020년 데이터를 활용하였으며 AUC를 평가지표로 하였다. 지역별로 각 모형의 예측성능을 평가한 뒤 혼동행렬을 집계하였으며, 전국 수준에서 예측성능을 평가하여 본 연구에서 사용한 방법론에 대해 고찰하도록 하였다. 통계분석 및 모델링을

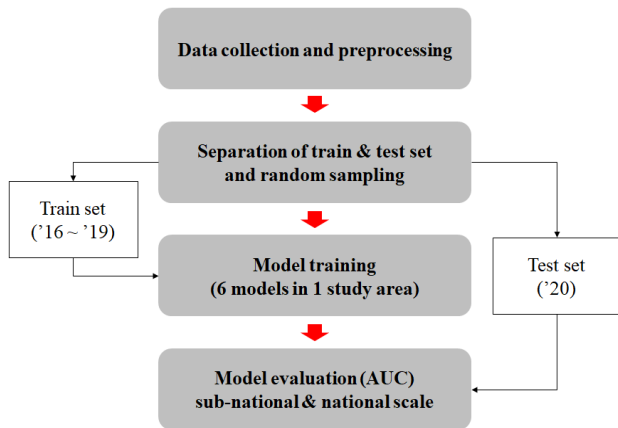


Fig. 5. Flowchart of Model Training and Model Testing

위해 Python 3.4 버전을 사용하였으며 sklearn, xgboost, imblearn 라이브러리를 통하여 샘플링 함수, 모델링 함수, 모델평가 함수, 그리고 오버샘플링 모형 중 하나인 SMOTE 함수를 호출하였다.

### 3.3 데이터 수집 및 가공 방법

모델학습에 사용된 데이터의 출처와 제원, 가공방법은 Table 2와 같다. 먼저, 기상정보에 해당하는 기온, 습도, 강수량, 풍속은 기상청의 기상자료개방포털을 활용하였고 사회경제정보에 해당하는 묘지 밀도와 농지 밀도는 통계청의 국가통계포털(KOSIS)을 활용하여 수집한 뒤 별도로 가공하였다. 산림정보에 해당하는 ha당 평균입목재적, 침엽수 비율, 사유림 비율은 산림청의 2020년 산림기본통계를 활용하였으며 시·군·구 단위에 맞게 가공하였다. 마지막으로, 본 연구에서 예측하고자 하는 산불발생여부는 산림청에서 전국을 대상으로 수집한 2016년부터 2020년까지의 자료를 활용하였다. 산림청에서 제공하는 원 데이터는 산불의 발생 일시, 진화일시, 발생장소, 발생원인 등의 정보를 포함하고 있으며 이를 전국의 시·군·구 지역을 기준으로 하여 일일 발생 횟수를 나타내는 데이터프레임 형태로 수정하여 활용하였다.

산불발생예측모형에 필요한 변수들은 별도로 가공이 필요하였는데 강수량의 경우 비가 내린 날을 포함하여 일정 기간 산불발생 예방효과가 있다고 가정하여, 날짜를 기준으로 하는 파생변수(변수명 : rain\_code)를 생성하였다. 풍속의 경우 전체 분포 그래프를 살펴보았을 때 이례적으로 풍속이 높은 날이 있다는 사실에 착안하여 이를 세 단계(매우 높음, 높음, 보통)로 구분하여 파생변수(변수명 : wind\_code)를 생성하였다. 마지막으로 산불발생여부는 1건 이상 발생한 경우를 1로 하고 한 번도 발생하지 않은 경우를 0으로 하는 이진형 변수(변수명 : fire\_occur)를 코드화하여 이를 종속변수로 활용하였다.

Table 2. List of Variables for Model Training

Types	Variables	Specifications
Meteorological status	Temperature	daily/SGG* temperature (°C)
	Humidity	daily/SGG humidity (%)
	Precipitation	daily/SGG precipitation (mm)
	Wind speed	daily/SGG wind speed (avg.) (m/s)
Socio-economic status	Cemetery density	SGG cemetery ratio (%)
	Farmland density	SGG farmland ratio (%)
Forest's status	Avg. Stem volume per ha	SGG stem volume per ha (avg.) (m <sup>3</sup> /ha)
	Conifer ratio	SGG conifer ratio (%)
	Private forest ratio	SGG private forest ratio (%)
Fire occurrence status		daily/SGG Frequency of fire occurrence

\*SGG : Si · Gun · Gu (municipal-level division of ROK)



### 3.4 샘플링 방법

산불이라는 사건은 발생하지 않은 날이 발생한 날보다 압도적으로 많으므로 산불발생 데이터는 대표적인 불균형 데이터라고 할 수 있다. 산림청 데이터에 따르면, 전국 8개 도에서 2016년부터 2020년까지 산불이 발생한 사건의 합계는 1,366건으로 산불이 발생하지 않은 사건의 합계인 66,556건의 2.1%밖에 되지 않았다. 지역별로 이러한 경향치를 확인한 결과, 전반적으로 1.0~3.4% 수준이었다(Fig. 6). 본 연구에서는 산불발생값의 불균형을 해소하기 위하여 학습데이터를 선정할 때 대표적인 오버샘플링 기법인 SMOTE (Synthetic Minority Over-sampling Technique)를 활용하여 산불발생 (fire\_occur = 1)과 미발생 (fire\_occur = 0)인 데이터가 1:1 비율이 되도록 하였다. SMOTE 기법은 오버샘플링을 진행할 때 데이터를 단순히 임의 추출하여 복제하지 않고 KNN (K-Nearest Neighbor) 알고리즘을 활용하여 동일 클래스 내에서 최대한 원본 데이터에 가깝도록 데이터를 증식하므로, 오버샘플링에서 단순 복제할 때 생길 수 있는 과적합 문제를 해소할 수 있어 유용하다.

### 3.5 모델학습 및 평가 방법

산불발생예측모형 생성을 위하여 LR (Logistic Regression) 과 SVM (Support Vector Machine), 그리고 앙상블 기법을 활용하였다. 앙상블 기법은 여러 개의 모형을 생성하고 그 예측값을 종합하여 최종적인 의사결정에 활용하는 방법으로 단일 모형을 활용할 경우보다 우수한 예측성능을 보인다. 본 연구에서 사용된 앙상블 기법은 Boosting 계열에 속하는 GB (Gradient Boosting)과 XGB (eXtreme Gradient Boosting), Bagging 계열에 속한 RF (Random Forest), 그리고 ET (Extra Tree)까지 총 4개였다. LR을 제외한 나머지 5개 기법의

경우 파이썬 라이브러리에서 종속변수가 범주형일 때 적용할 수 있는 분류기(classifier) 모델을 제공하므로 각 기법에 알맞은 분류기를 학습 및 예측에 사용하였으며, SVM의 경우 커널옵션을 'linear'로 설정하였다. 모델학습은 연구대상지별로 6개의 모형이 독립적으로 생성되도록 하였으며 이를 위하여 학습데이터를 지역별로 다시 분할하여 학습하도록 하였다.

모델학습에 사용된 분류기는 모델 예측값을 0~1 사이의 값으로 분류하며 이를 통해 모델 예측에 관한 성능을 평가할 수 있다. 예측모형의 성능 평가지표로는 정확도(accuracy), 민감도(sensitivity), 특이도(specificity) 등을 사용할 수 있으며 본 연구에서는 정확도 및 AUC (Area Under the ROC Curve)를 평가지표로 활용하였다. 정확도는 전체 입력된 데이터에 대해 얼마나 정확하게 예측했는가를 나타내는 지표로, 전체 데이터 건수 대비 예측값과 실제값이 동일한 건수로 계산한다. AUC는 1-특이도와 민감도를 각각 x, y축으로 하는 ROC (Receiver Operating Characteristic) 곡선의 아래 면적을 의미하며 다양한 임계값에서 모델의 분류 성능을 측정하는 데 사용된다. 본 연구에서는 대상지마다 가장 높은 AUC 값을 보여주는 모형을 선정하여 혼동행렬을 구성하고 8개 대상지의 혼동행렬을 결합하여 전국 수준 산불발생예측모형의 성능을 평가하였다.

## 4. 결과 및 고찰

본 연구에서는 산불예측모델의 평가지표를 정확도와 AUC로 설정하여 연구대상지별 6개 모형에 대한 예측성능을 평가하였다(Tables 3, 4). Site 1부터 Site 8까지 순서대로 경기도, 강원도, 충청북도, 충청남도, 전라북도, 전라남도,

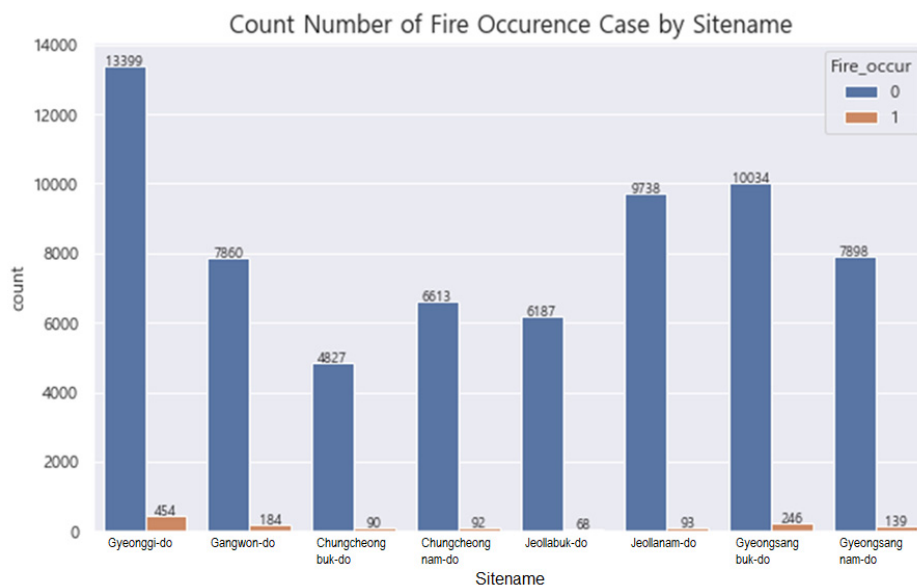


Fig. 6. Fire Occurrence by Province in 2016~2020

경상북도, 경상남도 지역을 의미하며, 산불발생예측모형의 유형은 순서대로 LR (Logistic Regression), SVC (Support Vector Classifier), GBC (Gradient Booster Classifier), XGBC (XGBoost Classifier), RFC (Random Forest Classifier), ETC (Extra Tree Classifier)이다. 정확도를 평가지표로 하였을 때 앙상블 기법을 이용한 모델이 모든 대상지에서 LR 또는 SVC보다 우수한 성능을 보여주었으며, 대상지에서 가장 우수한 모델은 RFC (Site 1, 2, 5, 6, 7, 8)와 GBC (Site 3, 4)였다. LR 및 SVC 모형의 경우 지역에 따라 산불 예측 정확도가 일정하지 않고 특히 강원도(Site 2) 지역에서는 현저하게 성능이 떨어지는 현상을 보였다. 각 지역에서 보이는 예측모델 정확도의 범위는 0.83~0.97 수준으로 나타났다. 한편, AUC를 평가지표로 하였을 때 성능이 우수한 모델은 앙상블 모델에만 국한되지 않았으며 LR과 SVC 모델은 Site 1, 5, 6, 7에서 우수한 성능을 보였으며, 앙상블 모형은 Site 2, 3, 4, 8에서 우수한 성능을 보여주었다. AUC를 평가지표로 하는 경우 각 모델의 성능에 큰 편차는 없었으며, 지역별 AUC 값의 범위는 0.64~0.81였다. 가장 예측성능이 떨어지는

지역은 충청북도(Site 3) 지역으로, 2020년 산불발생 건수가 11건으로 가장 적었던 것을 감안하면 산불발생과 관련된 정보의 부족으로 인해 예측이 어려웠을 것으로 보인다. Fig. 7은 연구대상지별 모형에 따른 ROC 곡선과 AUC 값을 나타낸 그래프로 위로 볼록한 곡선을 보여주고 있으며, 경기도와 경상남도(Site 1, 8) 지역의 경우 6개 모델별 ROC 곡선이 비슷한 형태를 보여주고 있다.

Table 5는 이상의 산불발생예측모형의 평가 결과에 따라 8개 지역별 혼동행렬을 결합한 전국 수준에서의 산불발생 예측 결과표다. 2020년 전국 8개 도에서 산불이 발생하지 않은 13,325건 중 8,290건이 발생하지 않았음을 예측(62.2%) 하였으며 산불이 발생한 333건 중 239건이 산불이 발생하였음을 예측(71.8%)하였다. 이를 종합한 산불발생예측모형의 정확도는 62.4%였다. 산불관리에 머신러닝을 적용한 논문들을 검토한 Bot and Borges (2022)에 의하면, 산불예측에 머신러닝 기법을 활용한 연구에서 AUC가 가장 높은 평가지표로 사용되고 있으며, 그 값은 0.68에서 0.99까지 다양하게 나타나고 있다. 그러나 대부분의 연구의 경우 산불발생의 위험성

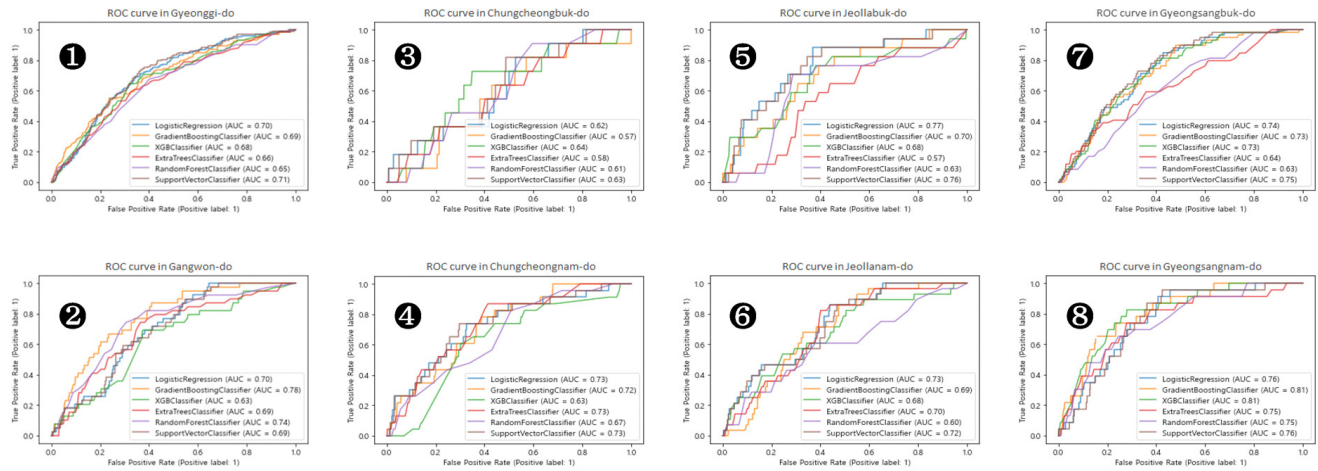
**Table 3.** Results of Evaluating Model Prediction according to 8 Research Sites (criterion : accuracy)

Types	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	Site 8
LR	0.507	0.119	0.649	0.694	0.641	0.462	0.638	0.427
SVC	0.528	0.109	0.718	0.579	0.629	0.407	0.453	0.611
GBC	0.812	0.941	0.897	0.983	0.886	0.761	0.737	0.708
XGBC	0.674	0.966	0.731	0.959	0.788	0.709	0.751	0.626
RFC	0.830	0.967	0.882	0.971	0.920	0.858	0.924	0.853
ETC	0.755	0.891	0.767	0.850	0.828	0.826	0.890	0.731
Model selection	<RFC> acc. = 0.83	<RFC> acc. = 0.97	<GBC> acc. = 0.90	<GBC> acc. = 0.98	<RFC> acc. = 0.92	<RFC> acc. = 0.86	<RFC> acc. = 0.92	<RFC> acc. = 0.85

**Table 4.** Results of Evaluating Model Prediction according to 8 Research Sites and Confusion Matrix (criterion : AUC)

Types	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	Site 8
LR	0.704	0.696	0.616	0.730	0.765	0.729	0.735	0.765
SVC	0.710	0.693	0.629	0.732	0.762	0.721	0.748	0.760
GBC	0.694	0.779	0.569	0.724	0.701	0.695	0.725	0.807
XGBC	0.676	0.631	0.639	0.626	0.684	0.681	0.731	0.808
RFC	0.648	0.739	0.612	0.666	0.627	0.599	0.625	0.749
ETC	0.656	0.692	0.584	0.733	0.566	0.697	0.636	0.752
Model selection	<SVC> AUC = 0.71	<GBC> AUC = 0.78	<XGBC> AUC = 0.64	<ETC> AUC = 0.73	<LR> AUC = 0.77	<LR> AUC = 0.73	<SVC> AUC = 0.75	<XGBC> AUC = 0.81
Observed number of forest fire occurred in 2020								
0: Non-occur	0: 2,653	0: 1,579	0: 979	0: 1,327	0: 1,240	0: 1,949	0: 2,010	0: 1,588
1: Occur	1: 133	1: 39	1: 11	1: 23	1: 17	1: 28	1: 59	1: 23
Confusion matrix*	[1361 1292] [ 24 109]	[1519 60] [ 36 3]	[719 260] [ 6 5]	[1138 189] [ 13 10]	[793 447] [ 4 13]	[889 1060] [ 3 25]	[882 1128] [ 4 55]	[989 599] [ 4 19]

\*This is a matrix with the point (0,0) as TN (true negative) and the point (1,1) as TP (true positive)



**Fig. 7.** ROC Curves for Each Model according to 8 Research Sites (From ① to ⑧ Shows Gyeonggi-do, Gangwon-do, Chungcheongbuk-do, Chungcheongnam-do, Jeollabuk-do, Jeollanam-do, Gyeongsangbuk-do, and Gyeongsangnam-do in Order)

**Table 5.** Model Evaluation at National Scale

Observed	Predicted		Accuracy
	Non-occur	Occur	
Non-occur	8,290	5,035	62.2%
Occur	94	239	71.8%
		Overall	62.4%

(risk)이나 산불 취약성(susceptibility) 등과 같이 실제값이 아닌 지표값을 예측하는 것에 만족하였으며, 산불 사건의 발생(occurrence)을 연구한 경우에는 AUC가 0.68~0.72 수준으로 나타났다. 국내에서는 Chae et al. (2018)의 연구에서 강원도 지역을 9개 구역으로 분할하여 머신러닝 기법과 오버샘플링을 적용한 사례가 있으며, 평가 지표는 정확도를 사용하였다. 강원도 지역만을 대상으로 하는 산불발생 예측 정확도는 0.76 수준으로 나타났다. 본 연구의 경우 8개 도를 대상으로 각 지역별 모델을 평가한 결과는 0.64~0.81 (AUC)이며, 전국수준 산불발생 예측 정확도는 약 0.72 (accuracy)로써, 기존의 연구보다 공간적·시간적 범위를 확장하였지만 대등한 수준의 예측 정확도가 나타났다.

본 연구에서 모델학습에 사용된 설명변수들이 산불발생에 미치는 영향을 평가하기 위해 LR모형에 대해서는 Table 6과 같이, 지역별로 생성된 모델 내 변수들의 계수의 값을 확인하였다. 각 지역에서 생성된 모델 내 설명변수의 계수는 일부 경우를 제외하고는 모두 p-value가 0.01 이하를 만족하고 있었다. 본 연구에서는 모든 지역에서 습도(HM)와 강수 코드(RC)가 산불발생에 있어서 중요한 설명변수로 나타났으며, 음(-)의 부호는 습도가 높아지거나 강수 코드가 증가할수록 산불발생을 억제하는 효과가 있음을 의미한다. 지역에 상관없이 기온(TP)이 높아질수록 산불발생 확률이 높아지며 대체로 풍속(WC)은 값이 커질수록 산불발생 확률이 낮아

**Table 6.** Coefficient Values Appearing in the LR Model

Site no.	Explanatory Variables*								
	TP	HM	RC	WC	CD	FD	SV	CR	PR
1	.419	-.684	-.411	-.297	.185	.193	.258	.323	-.111
2	.333	-.769	-.332	.086	.364	.951	.901	-.377	-.564
3	.365	-.951	-.666	-.531	-.106	-.269	-.116	.113	.333
4	.731	-.847	-1.028	-.084	.095	-.342	-.103	.243	.512
5	.722	-.950	-.590	-.029	-.402	-.162	.067	.188	-.224
6	.081	-.892	-.535	-.199	.333	.165	.537	.116	.156
7	.291	-.719	-.644	-.134	-.007	-.064	-.038	.147	-.021
8	.230	-.544	-.902	-.247	-.353	-.277	.174	-.160	.369

\*Abbreviation : TP (temperature), HM (humidity), RC (rain\_code), WC (wind\_code), CD (cemetery density), FD (farmland density), SV (stem volume per ha), CR (conifer ratio), PR (private forest ratio)

\*\*Cell color : (p-value > 0.01) (p-value > 0.1)

지는 것을 확인할 수 있었다. 기온은 봄철 외부활동의 증가와 관련이 있어 간접적인 요인으로 해석된다. 반면 풍속 변수의 경우 산불과 상관관계가 있지만 부호에서 일관성이 없었다. 이는 풍속의 경우 도나 시단위의 평균값 보다는 산지의 미세 지형에서 발생하는 순간 풍속에 직접 영향을 받기 때문에 기대하는 결과값이 나타나지 않은 것으로 보인다. 한편, 강원도(Site 2) 지역에서는 농지 밀도(FD)와 ha당 입목재적(SV)이 산불발생에 미치는 영향이 큰 것으로 나타났으며, 그 외의 지역에서는 산불발생에 미치는 영향이 불분명한 것으로 나타났다. 묘지 밀도(CD)의 경우, 전국 평균적으로 유의한 영향을 확인할 수 있었지만, 일부 지역(site 5, site 7)에서는 산불발생에 미치는 영향이 유의하지 않았다. 지형이나 인문 사회적 변수들은 날씨 변수와 달리 지엽적으로 영향을 미치



기 때문에 분석을 위한 지역의 단위(예, 읍·면·동 단위)를 줄이는 것이 모형의 예측력과 설명력을 높일 수 있을 것이다. 그 외의 요인들(CR, PR)의 경우 예측모델 안에서는 유의한 변수들이지만 지역별 모델별로 부호가 일정하지 않으므로 산불발생에 미치는 영향이 분명하지 않았다.

이상의 결과는, 산불에 영향을 미치는 변수들로 기온, 습도, 풍속 등의 기상변수를 주요 요인으로 제시한 Lee et al. (2004)와 Won et al. (2016)의 연구결과와 부합하는 한편, 습도의 경우 더욱 정밀한 산불발생 예측을 위해서 실효습도를 사용하는 방안이 제안된다. 강수량의 경우 산불발생건수와 밀접한 연관성을 가지는 것으로 알려져 있으며(Kim et al., 2013) 산불 억제에 미치는 영향 명확하여 산불관련 각종 지수에도 포함되나, 본 연구를 통하여 그 연관성을 산불발생 예측모형 내에서 실증적으로 규명하였다. 한편, Lee et al. (2011)은 산불발생확률에 영향을 주는 변수로써 산림특성을 나타내는 임분축적이나 수종 등 주요 인자를 고려할 것을 제안하였는데 본 연구에서는 ha당 입목축적, 소나무 비율, 사유림 비율과 같은 산림특성변수가 모두 유의한 변수로 나타났다. Lee et al. (2012)는 산불발생확률을 추정하기 위해 로지스틱 회귀모형을 사용하였으며, 묘지와 거리, 과거 산불빈도, 수종, 고도, 경사도 등이 산불발생에 영향을 주는 것으로 제안하였는데 본 연구에서는 묘지의 경우 모든 지역에서는 아니지만 산불발생예측 LR모형에서 유의한 변수로 나타났다. 이는 성묘객 등 산림이용빈도가 높은 지역에서 산불발생의 위험을 높아짐을 의미한다.

## 5. 결 론

기후변화에 따라 고온 건조한 날씨가 지속되면서 전 세계적으로 산불에 의한 피해가 증가하는 추세다. 국내에서도 2022년 3월에 발생한 울진-삼척 산불에 의한 피해면적이 20,923 ha에 달하였고, 건조한 기후의 영향으로 총 213시간 동안 산불이 이어지면서 역대 최장 산불로 기록되었다. 본 연구는 머신러닝 기법을 활용하여 산불기간 동안 전국의 지역별 산불발생을 예측하는 모형을 개발하는 것이다. 모델링 기법으로 로지스틱 회귀분석과 GB, RF 등의 앙상블 기법을 활용하였다. 또, 산불은 발생한 날보다 발생하지 않는 날이 많기 때문에 지역별 발생 패턴을 데이터로 나타내면 Null 값을 많이 가지게 되는데, 본 연구에서는 이를 해결하기 위해 오버샘플링 기법을 적용하였다. 개발한 모형을 이용하여 2020년 전국 산불기간(2월-4월) 산불발생을 예측한 결과 전체 333건 중에서 239건의 산불을 예측하였으며, 예측 정확도는 약 71.8%로 나타났다. 산불의 발생은 기온, 습도, 강수량 등 기후 요인에 가장 큰 영향을 받지만, 우리나라에서는 농지 밀도, 묘지 밀도, 사유림 비율 등 사회경제적 요인도 중요한 것으로 나타났다. 이는 국내에서 발생하는 대부분의 산불은 산림이나 인접지에서 사람들의 활동에

의해 발생하기 때문으로 해석된다.

본 연구에서는 전국을 지역별로 구분할 때, 광역도 단위에서 산불의 발생을 예측하는 연구를 개발하였다. 하지만 인구 사회학적인 요인이나 미세지형과 기후가 산불에 미치는 영향을 분석하기 위해서는 시·구·군 단위, 혹은 읍·면·동 단위에서 산불발생을 예측하는 모형을 개발하는 것이 필요하다. 향후 산불발생예측모형의 정밀도를 높이고 고도화하기 위해 다양한 공간데이터와 최신 리모트 센싱(remote sensing) 기법을 활용한 연구가 필요하다. 또한 산불의 예방과 초기 진화를 위한 진화자원 운영계획을 효율적으로 수립하기 위해 지역별 산불발생 요인을 규명하고 고도화된 산불 예측모형의 개발이 필요하다.

## 감사의 글

이 논문은 2022년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(NRF-2021R1F1A1046935).

## References

- Alonso-Betanzos, A., Fontenla-Romero, O., Guijarro-Berdiñas, B., Hernández-Pereira, E., Andrade, M.I.P., Jiménez, E., et al. (2003). An intelligent system for forest fire risk prediction and fire fighting management in Galicia. *Expert Systems With Applications*, Vol. 25, No. 4, pp. 545-554.
- Bot, K., and Borges, J.G. (2022). A systematic review of applications of machine learning techniques for wildfire management decision support. *Inventions*, Vol. 7, No. 1, 15.
- Chae, K.J., Lee, Y.R., and Park, J.H. (2018). Development of a Gangwon province forest fire prediction model using machine learning and sampling. *The Journal of Bigdata*, Vol. 3, No. 2, pp. 71-78.
- Cunningham, A.A., and Martell, D.L. (1973). A stochastic model for the occurrence of man-caused forest fires. *Canadian Journal of Forest Research*, Vol. 3, No. 2, pp. 282-287.
- Dutta, R., Aryal, J., Das, A., and Kirkpatrick, J.B. (2013). Deep cognitive imaging systems enable estimation of continental-scale fire incidence from climate data. *Scientific Reports*, Vol. 3, No. 1, pp. 1-4.
- Fried, J.S., Gillies, J.K., and Spero, J. (2006). Analysing initial attack on wildland fires using stochastic simulation. *International Journal of Wildland Fire*, Vol. 15, No. 1, pp. 137-146.
- Hankyoreh. (2022). *Uljin-Samcheok forest fire suppressed...*

- in 213 hours, "The longest record ever". Retrieved February 10, 2023, from [https:// www.hani.co.kr/arti/area/yeongnam/1034615.html](https://www.hani.co.kr/arti/area/yeongnam/1034615.html)
- Jain, P., Coogan, S.C., Subramanian, S.G., Crowley, M., Taylor, S., and Flannigan, M.D. (2020). A review of machine learning applications in wildfire science and management. *Environmental Reviews*, Vol. 28, No. 4, pp. 478-505.
- Kim, S.S., Lee, J.H., and Lee, M.W. (2013). Characteristics of forest fires and weathers in domestic over the past 50 years through the statistics. *Journal of the Korean Society of Hazard Mitigation*, Vol. 13, No. 5, pp. 225-232.
- Korea Forest Service (KFS). (2009). *Climate change and Forest*.
- Korea Forest Service (KFS). (2018). *Forest fire statistical yearbook*.
- Korea Forest Service (KFS). (2021). *Forest fire statistical yearbook*.
- Korea Forest Service (KFS). (2023). *Status of forest fire occurrence*. Retrieved February 10, 2023, from [https://www.forest.go.kr/kfsweb/kfi/kfs/frfr/selectFrfrStatsArea.do?mn=NKFS\\_02\\_02\\_01\\_05](https://www.forest.go.kr/kfsweb/kfi/kfs/frfr/selectFrfrStatsArea.do?mn=NKFS_02_02_01_05)
- Kwak, H.B., Lee, W.K., Lee, S.Y., Won, M.S., Koo, K.S., Lee, B.D., and Lee, M.B. (2010). Cause-specific spatial point pattern analysis of forest fire in Korea. *Journal of Korean Society of Forest Science*, Vol. 99, No. 3, pp. 259-266.
- Lee, B., Lee, Y., Lee, M.B., and Albers, H.J. (2011). Stochastic simulation model of fire occurrence in the Republic of Korea. *Journal of Korean Society of Forest Science*, Vol. 100, No. 1, pp. 70-78.
- Lee, B.D., Ryu, G.S., Kim, S.Y., and Kim, K.H. (2012). Development of forest fire occurrence probability model using logistic regression. *Journal of Korean Forestry Society*, Vol. 101, No. 1, pp. 1-6.
- Lee, H., Kim, Y., and Yoon, T. (2022). Estimation of environmental costs for deforestation: Focusing on the case of limestone mine development projects. *Journal of Environmental Policy and Administration*, Vol. 30, No. 1, pp. 63-91.
- Lee, S.Y., and Lee, H.P. (2006). Analysis of forest fire occurrence in Korea. *Fire Science and Engineering*, Vol. 20, No. 2, pp. 54-63.
- Lee, S.Y., Han, S.Y., Won, M.S., An, S.H., and Lee, M.B. (2004). Developing of forest fire occurrence probability model by using the meteorological characteristics in Korea. *Korean Journal of Agricultural and Forest Meteorology*, Vol. 6, No. 4, pp. 242-249.
- Lee, Y., Fried, J.S., Albers, H.J., and Haight, R.G. (2013). Deploying initial attack resources for wildfire suppression: Spatial coordination, budget constraints, and capacity constraints. *Canadian Journal of Forest Research*, Vol. 43, No. 1, pp. 56-65.
- Loepfe, L., Martinez-Vilalta, J., and Piñol, J. (2011). An integrative model of human-influenced fire regimes and landscape dynamics. *Environmental Modelling & Software*, Vol. 26, No. 8, pp. 1028-1040.
- Martell, D.L., Bevilacqua, E., and Stocks, B.J. (1989). Modelling seasonal variation in daily people-caused forest fire occurrence. *Canadian Journal of Forest Research*, Vol. 19, No. 12, pp. 1555-1563.
- Martell, D.L., Otukol, S., and Stocks, B.J. (1987). A logistic model for predicting daily people-caused forest fire occurrence in Ontario. *Canadian Journal of Forest Research*, Vol. 17, No. 5, pp. 394-401.
- Pechony, O., and Shindell, D.T. (2010). Driving forces of global wildfires over the past millennium and the forthcoming century. *Proceedings of the National Academy of Sciences*, Vol. 107, No. 45, pp. 19167-19170.
- Pew, K.L., and Larsen, C.P.S. (2001). GIS analysis of spatial and temporal patterns of human-caused wildfires in the temperate rain forest of Vancouver Island, Canada. *Forest Ecology and Management*, Vol. 140, No. 1, pp. 1-18.
- Rodrigues, M., and De la Riva, J. (2014). An insight into machine-learning algorithms to model human-caused wildfire occurrence. *Environmental Modelling & Software*, Vol. 57, pp. 192-201.
- Sakr, G.E., Elhajj, I.H., Mitri, G., and Wejinya, U.C. (2010). Artificial intelligence for forest fire prediction. In *2010 IEEE/ASME international conference on advanced intelligent mechatronics* (pp. 1311-1316). IEEE.
- Stojanova, D., Kobler, A., Ogrinc, P., Ženko, B., and Džeroski, S. (2012). Estimating the risk of fire outbreaks in the natural environment. *Data Mining and Knowledge Discovery*, Vol. 24, pp. 411-442.
- Todd, J.B., and Kourtz, P.H. (1991). *Predicting the daily occurrence of people-caused forest fires*. Petawawa National Forestry Institute, Information Report PI-X-103. Natural Resources Canada, Ottawa, Ontario, Canada.
- Van Beusekom, A.E., Gould, W.A., Monmany, A.C., Khalyani, A.H., Quiñones, M., Fain, S.J., et al. (2018). Fire weather and likelihood: Characterizing climate

- space for fire occurrence and extent in Puerto Rico. *Climatic Change*, Vol. 146, pp. 117-131.
- Vasilakos, C., Kalabokidis, K., Hatzopoulos, J., and Matsinos, I. (2009). Identifying wildland fire ignition factors through sensitivity analysis of a neural network. *Natural Hazards*, Vol. 50, pp. 125-143.
- Vecín-Arias, D., Castedo-Dorado, F., Ordóñez, C., and Rodríguez-Pérez, J.R. (2016). Biophysical and lightning characteristics drive lightning-induced fire occurrence in the central plateau of the Iberian Peninsula. *Agricultural and Forest Meteorology*, Vol. 225, pp. 36-47.
- Won, M., Yoon, S., and Jang, K. (2016). Developing Korean forest fire occurrence probability model reflecting climate change in the spring of 2000s. *Korean Journal of Agricultural and Forest Meteorology*, Vol. 18, No. 4, pp. 199-207.

---

<b>Received</b>	March 13, 2023
<b>Revised</b>	March 14, 2023
<b>Accepted</b>	March 20, 2023