

# 주택화재에서 주거지 유형별 재산피해에 영향을 미치는 화재인자 영향분석에 관한 연구

정원일  
호서대학교 컴퓨터공학부

## A Study on the Impact Analysis of Fire Factors Affecting Property Damage by Residential Type in House Fires

Weonil Jeong  
Division of Computer Engineering, Hoseo University

**요약** 최근 주거지 건축물이 고층화되거나 과밀화되는 등 다양한 형태의 변화 요인으로 인해 잠재적인 화재 발생의 위험성도 높아지고 있다. 2020년 소방청 통계연보에 따르면 과거 10년간 연평균 주택화재 건수는 0.6%, 인명피해와 재산피해도 각각 1.6%, 2.4%씩 증가하고 있다. 이에 본 논문에서는 주택화재로 인한 재산피해를 최소화할 수 있도록 빅데이터 기반의 데이터 분석을 통해 주택화재 데이터로부터 공동주택, 단독주택, 기타주택 등 주거지 유형에 따라 재산 피해에 영향을 미치는 화재요인들을 평가하고, 도출 결과로부터 주택화재에 대한 화재 예방 및 대응을 위한 제도적 개선 방안을 제시한다. 제안 논문에서 화재인자 영향분석은 유의수준을 만족하는 화재 인자를 선정하는 유의확률 분석, 재산 피해와 선정 화재 인자 간의 선형적 상관관계 분석, 랜덤포레스트 알고리즘을 이용하여 주거지 유형별 재산피해에 영향을 미치는 화재 인자들의 중요도를 도출하는 과정으로 수행된다. 주요 결과로 주택화재에 영향을 미치는 공통적인 중요 화재 인자는 소방서거리, 출동소요시간, 화재진압시간 등으로 분석되었으며, 이러한 결과에 대한 제도적 개선 방안으로 소방서 위치 최적화, 출동 지연 요인 감소 방안 등이 제고되어야 한다. 또한, 고층아파트나 주상복합건물이 포함된 공동주택의 경우 다른 주거지 유형에 비해 발화층이 상대적으로 높은 중요도를 보였는데, 이를 위한 개선 방안으로는 고층 화재 예방과 대응을 위해 소화설비 설치 강화 등의 노력이 필요함을 알 수 있었다.

**Abstract** Recently, the potential risk from residential building fire is increasing due to various factors such as high-rise or overcrowding. According to the 2020 Korean National Fire Agency Statistical Yearbook, the average annual number of housing fires over the past ten years has increased by 0.6%, and the number of casualties and property damage have increased by 1.6% and 2.4%, respectively. Therefore, this paper evaluates the fire factors that affect property damage according to the type of residence, such as multi-unit houses, detached houses, and other houses. In particular, this study used housing fire data for big data analysis to minimize property damage caused by house fires. Moreover, this study proposes institutional improvement plans for fire prevention and response to house fires from the derived results. This paper's fire factor impact analysis includes a significance probability analysis to select a fire factor that satisfies the significance level. In addition, a linear correlation analysis was performed between property damage and the selected fire factor, and a random forest algorithm was implemented as a process of deriving the importance of factors. As the main results, the common important fire factors affecting house fires were fire station distance, time required for dispatch, and fire suppression time. As a systematic improvement plan for these results, it is necessary to improve the location optimization of the fire station and reduce the delay factors for dispatch. In addition, in the case of multi-unit houses, including high-rise apartments or multi-purpose buildings, the fire floor showed relatively high importance compared to other residential types.

**Keywords** : House Fires, Property Damage, Fire Factors, Impact Analysis, Random Forest

---

\*Corresponding Author : Weonil Jeong(Hoseo Univ.)

email: wncung@hoseo.edu

Received October 12, 2021

Accepted February 4, 2022

Revised November 4, 2021

Published February 28, 2022

## 1. 서론

최근 주거지 건축물은 고층화되고 과밀화되는 추세를 나타내고 있다. 이러한 주거지에서 다양한 잠재적 화재 요인으로 인한 화재 발생의 위험성이 증가하고 있다. 주거지에서의 화재사고는 다양한 유무형의 사회적인 손실로 이어지기 때문에 주거지 화재 안전에 관한 사회적 관심도 높아지고 있다[1,2]. 2020년 소방청 통계연보에 따르면 최근 10년간 주거지에서 발생한 화재는 연평균 11,125건으로 0.6% 증가세를 보였다. 또한, 동기간 인명피해와 재산피해도 연평균 1.6%, 2.4%씩 늘어나고 있다[3].

주거지 화재에서 사회적 손실을 최소화하기 위해 공동주택의 화재 취약요인에 대해 화재안전성능 인증제도를 제안한 연구, 공동주택 화재안전규정에 관한 체크리스트에 대해 중요도와 검증 수준을 분석하여 효율적인 화재 안전규정 매뉴얼 개발에 관한 연구, 그리고 고층 아파트의 피난규정 현황을 분석하여 고층 아파트 피난기준의 문제점 도출 및 피난 안전성 강화를 위한 제도적인 자료 활용 방안에 관한 연구들이 수행되었다[4-6]. 그러나 효율적인 주거지 화재 예방 및 대응을 위해서는 화재에 내재된 여러 화재 인자들의 상관성을 기반으로 하는 복합적인 위험 요인의 분석이 전제되어야 한다. 이에 최근 소방청에서 제공하는 공공데이터를 활용하여 빅데이터 기반의 기계학습 방법론을 기반으로 화재 발생에 따른 재산피해액과 화재 발생 여부를 예측하여 건물별 화재 위험도에 관한 연구[7], 화재에서 부상이나 사망 등 인명피해에 영향을 미치는 화재요인에 대한 영향평가 및 분석에 관한 연구[8], 그리고 소방청 화재정보에서 모든 화재현장을 대상으로 재산피해를 야기하는 화재요인을 분석하여 화재예방 강화를 위한 제도적 개선 방안이 연구되었다[9].

본 논문에서는 소방청에서 제공하는 화재 데이터에서 주택화재를 대상으로 주거지 유형에 따라 재산피해에 영향을 미치는 주요한 화재요인들을 평가하고 분석하여 주택화재에서의 화재에 관한 빅데이터 기반의 정책적 대안을 제시하고자 한다. 제안 논문에서는 소방청 공공데이터인 「소방청\_화재발생 주소정보」[10]에서 전체 183개 칼럼 가운데 빅데이터 분석을 위해 재산피해를 포함하여 수치화된 33개 칼럼을 대상 인자로 선정하여 실험하였다. 해당 데이터에서 주택은 공동주택, 단독주택, 기타주택으로 주거지 유형을 분류하고 있다. 주택화재에서 주거지 유형별 재산피해 영향분석은 먼저 화재 인자별 최

소값, 1사분위값, 중위값, 중간값, 3사분위값, 최대값 등의 기술 통계량을 분석한다. 이어 재산피해와 화재 인자 간의 유의확률을 평가하여 영향분석 과정에 유의미한 화재 인자를 선정한다. 그리고 재산피해와 선정된 화재 인자 간의 선형적 상관관계를 분석한다. 화재인자 영향평가 과정은 랜덤포레스트 알고리즘의 주요 파라미터 최적화, 화재 인자 값을 동일 모델 계수에 따른 가중치 부여를 위한 표준화, 재산피해와 분석 대상인 화재 인자의 영향예측 모델 생성 및 검증, 산출된 영향도에 관한 효과적인 해석을 위한 상대적 중요도를 제시하는 과정으로 수행하였다. 실험 결과에서는 공동주택에서 소방서거리 인자가 가장 높은 중요도를 보인 반면 단독주택이나 기타주택에서는 출동소요시간 인자가 가장 높은 결과를 보였다. 이로부터 공동주택 화재에서는 화재조기진압을 통해 재산피해 최소화를 위해 출동소요시간을 줄이기 위한 법과 제도적인 정비가 필요하다는 것을 알 수 있었다. 발화층 인자의 경우 공동주택(73.4%), 단독주택(19.3%), 기타주택(2.4%)의 순서로 나타났다. 이로부터 공동주택의 하나인 고층아파트에 대한 구조적인 특성을 고려하여 고층에서의 화재 예방 및 대응을 위해 소방설비 설치 강화 등의 노력이 필요하다는 주요 결과를 도출하였다. 제안 논문은 주택화재로 인한 재산피해를 감소시키기 위한 제도적으로 화재 예방 및 안전대책의 강화할 수 있도록 과학적 근거를 제시하고자 하였다.

본 논문의 구성은 다음과 같다. 2장에서 관련 연구에 관해 기술하고, 3장에서는 주거지 유형별로 재산피해에 영향을 미치는 인자에 관한 영향도 평가 및 분석과정을 논의하고, 4장에서 결론을 맺고 향후 연구를 제시하였다.

## 2. 관련 연구

### 2.1 유의확률

유의확률은 통계적인 가설검증에서 귀무가설이 참일 때 도출된 결과에 비해 실측 결과가 극단적으로 나타날 확률을 의미한다[11,12]. 유의확률은 0과 1 사이의 수치로 표현되며, 0에 수렴할수록 신뢰도가 높고 1에 수렴할수록 도출 결과보다 극단적 실측 결과가 발생할 확률이 높다는 것을 의미한다. 본 연구에서는 재산피해 화재 인자와 비교 대상인 화재 인자들 사이에 유의확률 산출하여 제안적 증거로 표현되는 유의수준을 만족하는 화재 인자에 대해 재산피해에 미치는 화재인자 영향평가를 실시한다. 이때 유의수준은 분석과정에서 임의 설정이 가

능하며 실험에서는 통상적인 유의수준으로 0.05를 적용한다.

## 2.2 상관관계 분석

상관관계는 한 변수의 변화에 따라 다른 변수가 연동하여 변화하는 정도를 나타낸다. 본 연구에서는 피어슨 상관관계 분석 방법[13,14]을 이용하여 연속하는 재산피해 변수와 분석 대상의 변수 사이의 선형적인 관계를 수치로 계량화한 피어슨 상관계수( $r$ )를 도출한다. 두 변수의 공분산을 각각의 표준편차의 곱으로 나눈 값으로 계산되는 피어슨 상관계수( $r$ )는 두 변수의 선형적인 상관관계에 따라 -1에서 1 사이의 값을 가진다. 이때 상관계수는 1에 수렴할수록 두 변수 간에 높은 선형적 상관관계가 존재하고, 0에 수렴할수록 선형적 상관관계가 낮다는 것을 의미하며, 만약 상관계수가 0.1 이하이면 두 변수 간의 선형적 상관관계는 무의미하다는 것을 의미한다.

## 2.3 랜덤 포레스트 기법

랜덤 포레스트 기법[15-17]은 검출, 분류, 회귀 분석 등에 사용되는 앙상블 학습방법이다. 해당 기법에서는 다수의 결정 트리를 앙상블 형태로 결합하고 있어 신속하고 정확한 훈련 성능과 대량의 데이터 처리를 지원한다. 랜덤 포레스트 수행 과정은 훈련 데이터를 입력으로 전체 특성(feature) 가운데 임의의 표본을 추출하고, 포레스트(forest)를 구성하는 개별 결정 트리(decision tree)의 예측값이 비상관화(decorrelation)되는 방향으로 다수의 서로 다른 결정 트리를 생성한다. 그리고 개별 결정 트리에서 최종적인 예측값이 산출될 때까지 분기를 반복하는 하향식 훈련을 수행한다. 이때 개별 결정 트리의 예측 결과는 과반수 투표제 방식으로 랜덤 포레스트를 결합하여 변수의 중요도를 최종적인 결과로 도출한다. 랜덤포레스트 기법은 앙상블 구조로 다수의 결정 트리가 결합되어 훈련 성능이 정확하고 대용량 데이터 처리를 지원한다. 본 논문에서는 랜덤포레스트 알고리즘을 활용하여 재산피해에 관한 화재 인자들의 영향평가 분석에 활용한다.

# 3. 본론

## 3.1 분석 데이터

소방청에서 제공하는 공공 화재 데이터[10]의 「장소중

분류」 속성에서 주택은 주거지 유형에 따라 공동주택(Multi-unit house), 단독주택(Detached house), 기타주택(Etc house)으로 분류하고 있다. 공동주택에는 기숙사, 다세대주택, 아파트, 연립주택, 주상복합아파트, 기타공동주택 등이 포함되며, 단독주택에는 다가구주택, 다중주택, 단독주택, 상가주택, 기타단독주택으로 나뉘며, 기타주택은 컨테이너, 비닐하우스 등으로 구분된다.

2017년도에 발생한 전체 화재 44,178건에 비해 주택에서 발생한 화재는 11,765건으로 26.6%의 비중을 차지한다. Fig. 1은 2017년도에 발생한 주택화재 건수와 비율을 공동주택, 단독주택, 기타주택으로 구분하여 분석한 그래프이다.

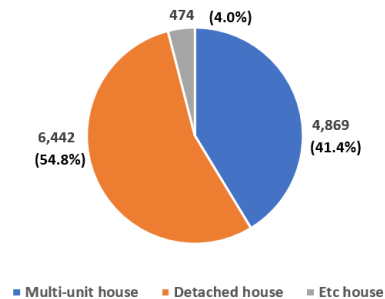


Fig. 1. Status of house fires in 2017

Fig. 1에서 2017년 발생한 주택화재는 공동주택 화재 6,442건(54.8%), 단독주택 4,869건(41.4%), 기타주택 474건(4.0%)의 순으로 나타났다.

소방청 공공 화재 데이터에서 개별화재에 관한 정보는 조사서번호, 사망, 부상, 인명피해(명), 재산피해소계, 시도본부, 소방서, 화재발생년월일 등 총 183개의 칼럼으로 구성이 가운데 주택화재에서 주거지 유형별 재산피해에 관한 화재인자의 영향분석에는 유의확률 및 상관관계 분석, 그리고 화재영향 평가를 위해 수치화된 칼럼을 대상으로 한다. 분석에 사용되는 칼럼은 총 33개의 칼럼으로, 종속 변수인 재산피해소계(PropertyDamage)를 포함하여 독립 변수로 사용되는 사망(Death), 부상(Injury), 인명피해소계(Human Damage), 출동소요시간(Arrival Time), 화재초진시간(Initial Extinguish Time), 화재진압시간(Total Extinguish Time), 건물층수\_지상(Ground Floor Level), 건물층수\_지하(Basement Level), 발화층(Fire Outbreak Level), 소방서거리(FireStation Distancing), 안전센터거리(SafeCenter Distancing), 소실면적(Burnt Area), 동원인력소계(Mobilized Personnel),

소방동원명수(Firefighting Personnel), 의소대동원명수(Volunteer Personnel), 경찰동원명수(Police Personnel), 일반직동원명수(Public Personnel), 군인동원명수(Military Personnel), 전기가스유관기관동원명수(ElectroGas Personnel), 기타동원명수(Etc Personnel), 동원장비소계(Mobilized Equipment), 펌프물탱크(PumpTank Equipment), 고가\_굴절(Aerial Equipment), 화학(Chemical Equipment), 구조(Rescue Equipment), 구급(Emergency Equipment), 헬기(Helicopter Equipment), 온도(Temperature), 습도(Humidity), 건물구조동(Building Block), 연면적(Total Floor Area), 바닥면적(Floor Area)이 해당된다.

상기 변수들의 통계 정보 분석을 위해 먼저 결측값(missing value)은 극단값(extreme outliers)에 대한 민감도가 낮은 중간값(median value)를 적용하여 설정하였다. Table 1에서는 분석에 이용하는 변수들의 최소값(Min), 1사분위수(First Qu.), 중간값(Meidan), 평균값(Mean), 3사분위수(Third Qu.), 최대값(Max)의 통계량 요약정보를 나타낸다.

Table 1에서 재산피해소계(PropertyDamage), 출동 소요시간(Arrival Time), 화재초진시간(Initial Extinguish Time), 화재진압시간(Total Extinguish Time), 건물층수\_지상(Ground Floor Level), 건물층수\_지하(Basement Level), 발화층(Fire Outbreak Level), 소방서거리(FireStation Distancing), 안전센터거리(SafeCenter Distancing), 소실면적(Burnt Area), 동원인력소계(Mobilized Personnel), 소방동원명수(Firefighting Personnel), 경찰동원명수(Police Personnel), 펌프물탱크(PumpTank Equipment), 고가\_굴절(Aerial Equipment), 연면적(Total Floor Area), 바닥면적(Floor Area)은 중간값이 1사분위수에 비교적 치우쳐진 분포를 보였다. 동원장비소계(Mobilized Equipment), 구조(Rescue Equipment), 구급(Emergency Equipment), 온도(Temperature), 습도(Humidity), 건물구조동(Building Block)의 경우에는 대체로 중간값이 평균값에 수렴하는 분포를 나타내었으며, 사망(Death), 부상(Injury), 인명피해소계(Human Damage), 의소대동원명수(Volunteer Personnel), 일반직동원명수(Public Personnel), 군인동원명수(Military Personnel), 전기가스유관기관동원명수(ElectroGas Personnel), 기타동원명수(Etc Personnel), 화학(Chemical Equipment), 헬기(Helicopter Equipment)는 최솟값, 1사분위수, 중간값, 3사분위수가 모두 0으로 극단적인 치우침을 나타내었다.

Table 1. Variables Summary

Fire factor	Min	First Qu.	Median	Mean	Third Qu.	Max
Death	0	0	0	0.02	0	3.0E+00
Injury	0	0	0	0.1	0	1.0E+01
HumanDamage	0	0	0	0.1	0	1.0E+01
PropertyDamage	0	75	496	5073	3690	7.5E+05
ArrivalTime	0	203	273	343	389	1.3E+04
InitialExtinguishTime	0	60	240	447	540	3.0E+04
TotalExtinguishTime	0	67	360	913	998	3.0E+04
GroundFloorLevel	0	1	3	6	6	5.9E+01
BasementLevel	0	0	0	0.5	1	1.2E+01
FireOutbreakLevel	-5	1	1	3	3	4.1E+01
FireStationDistancing	0	3	5	9	11	9.9E+01
SafeCenterDistancing	0	1	2	4	4	9.9E+01
BurntArea	0	0	0	16	15	1.5E+03
MobilizedPersonnel	0	22	32	36	48	1.1E+03
FirefightingPersonnel	0	16	24	29	41	2.3E+02
VolunteerPersonnel	0	0	0	2	0	3.0E+02
PolicePersonnel	0	2	2	3	4	3.5E+01
PublicPersonnel	0	0	0	0	0	1.2E+02
MilitaryPersonnel	0	0	0	0	0	9.7E+02
ElectroGasPersonnel	0	0	0	1	0	4.0E+01
EtcPersonnel	0	0	0	1	0	8.4E+01
MobilizedEquipment	0	8	11	11	15	5.0E+01
PumpTankEquipment	0	4	5	5.6	8	2.8E+01
AerialEquipment	0	0	0	0.5	1	1.0E+01
ChemicalEquipment	0	0	0	0.1	0	3.0E+00
RescueEquipment	0	1	1	1.3	2	1.0E+01
EmergencyEquipment	0	1	1	1.3	2	1.4E+01
HelicopterEquipment	0	0	0	0	0	4.0E+00
Temperature	-18	2	12	12	22	7.9E+01
Humidity	0	34	55	53	74	9.6E+02
BuildingBlock	0	1	1	2	1	3.1E+02
TotalFloorArea	0	90	276	44503	1931	2.1E+08
FloorArea	0	66	105	5494	270	9.7E+06

### 3.2 화재인자 선정

주책화재에서 재산피해에 관한 화재인자 영향을 분석하기 위해 먼저 종속 변수인 재산피해와 독립 변수인 화재인자 간의 유의확률(P-value)을 산출한다. 이어 유의수준을 만족하는 독립 변수를 선정하고, 재산피해와 화재인자 간의 피어슨 상관관계를 분석하여 선형적 상관 정도를 분석하였다.

종속 변수인 재산피해와 독립 변수인 화재인자 간의 유의확률(P-value)을 분석한 결과는 Table 2와 같다. Table 2의 유의확률 분석결과에서 유의수준이 0.05보다

작은 독립 변수는 유의수준을 만족하므로 이후 화재인자 분석 대상으로 선정한다. 소방동원명수(Firefighting Personnel), 군인동원명수(Military Personnel), 습도(Humidity), 건물구조동(Building Block), 연면적(Total Floor Area), 바닥면적(Floor Area)은 유의수준을 만족하지 않으므로 피어슨 상관관계 분석 및 화재인자 분석 대상에서 제외하였다.

Table 2. Analysis results of P-value

Fire factor	P-value
Death	<0.0000000000000002
Injury	<0.0000000000000002
HumanDamage	<0.0000000000000002
ArrivalTime	<0.0000000000000002
InitialExtinguishTime	<0.0000000000000002
TotalExtinguishTime	<0.0000000000000002
GroundFloorLevel	<0.0000000000000002
BasementLevel	<0.0000000000000002
FireOutbreakLevel	0.00000000004
FireStationDistancing	<0.0000000000000002
SafeCenterDistancing	<0.0000000000000002
BurntArea	<0.0000000000000002
MobilizedPersonnel	<0.0000000000000002
FirefightingPersonnel	0.8
VolunteerPersonnel	<0.0000000000000002
PolicePersonnel	<0.0000000000000002
PublicPersonnel	<0.0000000000000002
MilitaryPersonnel	0.3
ElectroGasPersonnel	<0.0000000000000002
EtcPersonnel	<0.0000000000000002
MobilizedEquipment	<0.0000000000000002
PumpTankEquipment	<0.0000000000000002
AerialEquipment	0.000000000000006
ChemicalEquipment	<0.0000000000000002
RescueEquipment	0.007
EmergencyEquipment	<0.0000000000000002
HelicopterEquipment	0.000000000000006
Temperature	0.000009
Humidity	0.09
BuildingBlock	0.6
TotalFloorArea	0.7
FloorArea	0.5

재산피해와 독립 변수 간의 선형적 상관관계의 정도를 분석하기 위한 피어슨 상관계수(Cor) 분석결과는 Table 3과 같다. Table 3에서 재산피해(Property Damage)에 대해 화재초진시간(Initial Extinguish Time), 화재진압 시간(Total Extinguish Time), 소실면적(Burnt Area), 기타동원명수(Etc Personnel), 의소대동원명수(Volunteer Personnel), 경찰동원명수(Police Personnel)가 상대적으로 높은 상관관계를 보임을 알 수 있다.

Table 3. Analysis results of Cor

Fire factor	Cor
Death	0.13
Injury	0.15
HumanDamage	0.18
ArrivalTime	0.11
InitialExtinguishTime	0.29
TotalExtinguishTime	0.35
GroundFloorLevel	-0.089
BasementLevel	-0.091
FireOutbreakLevel	-0.061
FireStationDistancing	0.13
SafeCenterDistancing	0.14
BurntArea	0.65
MobilizedPersonnel	0.14
VolunteerPersonnel	0.23
PolicePersonnel	0.22
PublicPersonnel	0.1
ElectroGasPersonnel	0.17
EtcPersonnel	0.26
MobilizedEquipment	0.13
PumpTankEquipment	0.093
AerialEquipment	-0.069
ChemicalEquipment	0.13
RescueEquipment	0.025
EmergencyEquipment	0.11
HelicopterEquipment	0.063
Temperature	-0.041

### 3.3 화재인자 영향분석

중속 변수인 재산피해에 대해 독립 변수인 화재 인자들의 영향도를 평가하고 분석하기 위해 사용한 실험 환경은 MS Windows 10 64비트 운영체제, Intel Xeon 3.5Ghz cpu, 128GB로 구성된 워크스테이션에서 RStudio를 사용하였다.

중속 변수인 재산피해는 학습 알고리즘에 영향이 없으므로 데이터 정제 대상에서 제외하였다. 독립 변수는 변수의 값이 모델 계수와 같은 가중치가 부여되도록 값을 조정하는 표준화를 수행하였다. 독립 변수의 표준화는 변수값에서 평균값을 차감하는 중앙화를 진행하였으며, 이 값을 해당 변수의 표준편차로 나누는 스케일링을 적용하였다.

Fig. 2는 스케일링을 수행하기 이전 단계의 독립 변수값 일부부를 예시로 나타내고 있다. Fig. 3은 스케일링을 수행한 이후 단계에서의 독립 변수값을 예시로 나타내고 있다.

데이터 정제를 표준화 단계인 스케일링을 적용한 독립 변수 데이터로부터 화재인자 영향분석을 위해 랜덤포레스트 알고리즘을 적용하였다. 알고리즘에서는 입력 데이터를 80%의 훈련 데이터와 20%의 검증 데이터로 분리하여 화재 인자의 영향 정도를 파악하기 위한 예측 모델을 생성하였다.

Injury	HumanDamage	ArrivalTime	InitialExtinguishTime
<dbl>	<dbl>	<dbl>	<dbl>
1	2	221	0
2	3	339	480
2	3	295	1140
2	3	608	780
1	2	264	1020

Fig. 2. Sample of variable's value list before scaling

Injury	HumanDamage	ArrivalTime	InitialExtinguishTime
-1.10	-1.10	-0.813	-1.49
0.73	0.73	-0.042	-0.45
0.73	0.73	-0.329	1.00
0.73	0.73	1.716	0.21
-1.10	-1.10	-0.532	0.73

Fig. 3. Sample of variable's value list after scaling

랜덤포레스트 알고리즘의 실행 결과에 대한 신뢰성 향상을 위해 랜덤포레스트 알고리즘에 이용되는 파라미터의 최적화가 요구된다. 이에 실험에서는 재산피해 및 독립 변수들을 포함한 학습 데이터에 대해 stepFactor는 0.5, improve는 1e-5를 입력으로 tuneRF 함수를 이용하여 최적화 과정을 수행하였다. Fig. 4는 랜덤포레스트 알고리즘의 주요 파라미터인 mtry 값의 변화에 따른 OOB 오차(out-of-bag error)의 산출한 결과를 나타낸다.

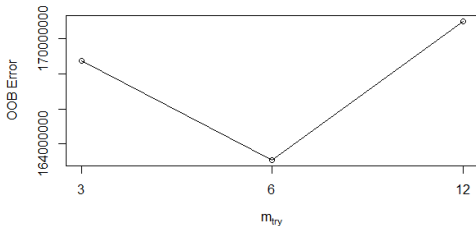


Fig. 4. OOB error progress of mtry

Fig. 4의 결과로부터 랜덤포레스트 알고리즘의 주요 파라미터인 mtry가 6일 때 OOB오차가 가장 낮은 16307635로 산출되었다. 이에 mtry의 최적값은 6으로 결정하였으며, ntree는 500을 적용하였다.

화재인자의 영향분석 과정은 최적 파라미터값과 재산피해 및 화재 인자 훈련 데이터를 입력으로 회귀 형식의 랜덤포레스트 예측 모델을 생성한다. 예측 모델의 성능은 실제값과 예측값의 편차를 나타내는 평균제곱근 오차를 나타내는 RMSE(Root Mean Square Error)를 적용하였다. 예측 모델을 활용한 실험 결과에서 독립 변수의 중요도 선정을 위해 산출된 %IncMSE(정확도), IncNodePurity(중요도)를 이용한다. 그리고 최종 중요도는 varImpPlot 함수로 시각화하고 IncNodePurity를 기준으로 도출하였다.

주택화재에서 재산피해에 미치는 화재인자의 영향 정도를 분석하는 과정은 주거지 유형으로 분류한 공동주택, 단독주택, 기타주택에 대한 각각의 예측 모델 실험 결과를 제시하고, 주거지 유형을 종합하여 분석하는 과정으로 설명한다.

공동주택(Multi-unit house)의 화재에서 재산피해에 영향을 미치는 화재 인자의 랜덤포레스트 분석결과는 Table 4와 같다. 이때 예측 모델의 평균제곱근오차는 18,222로 나타났다.

Table 4. Analysis result: Multi-unit house fires

Fire factor	%IncMSE	IncNodePurity
Death	8.28E+05	2.71E+09
Injury	7.89E+04	2.52E+09
HumanDamage	1.05E+06	2.83E+09
ArrivalTime	6.65E+06	1.02E+11
InitialExtinguishTime	5.89E+07	5.78E+10
TotalExtinguishTime	6.85E+07	9.6E+10
GroundFloorLevel	1.00E+07	5.31E+10
BasementLevel	2.48E+06	1.12E+10
FireOutbreakLevel	1.95E+07	9.25E+10
FireStationDistancing	-2.15E+06	1.26E+11
SafeCenterDistancing	5.57E+06	7.14E+10
BurntArea	1.12E+07	5.35E+10
MobilizedPersonnel	7.34E+07	6.98E+10
PolicePersonnel	8.86E+05	6.58E+10
MobilizedEquipment	5.87E+07	5.02E+10
PumpTankEquipment	4.96E+07	4.07E+10
AerialEquipment	1.36E+07	1.37E+10
RescueEquipment	1.63E+07	1.47E+10
EmergencyEquipment	4.40E+06	1.43E+10
Temperature	-2.41E+06	8.25E+10

Table 4는 공동주택 화재에서 재산피해에 대해 화재인자들의 영향 정도를 나타내는 정확도와 중요도를 확인할 수 있으며, 중요도 값이 클수록 재산피해에 미치는 영향이 증가함을 의미한다.

Fig. 5는 Table 4의 분석결과로부터 공동주택 화재에서 재산피해에 대한 독립 변수인 화재 인자의 중요도를 시각화하였다.

Fig. 5에서 공동주택 화재에서 재산피해에 대한 화재인자 중요도는 소방서거리, 출동소요시간, 화재진압시간, 발화층, 온도, 안전센터거리, 동원인력의 순으로 나타났다.

단독주택(Detached house)에서의 화재로 인한 재산피해에 영향을 미치는 인자들의 영향에 관한 랜덤포레스트 분석결과는 Table 5와 같이 산출되었으며, 예측 모델의 평균제곱근오차는 17,318이다.

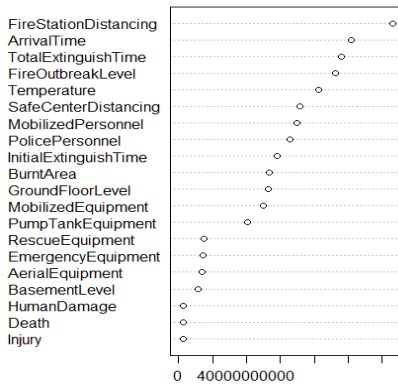


Fig. 5. Importance of independent variables: Multi-unit house fires

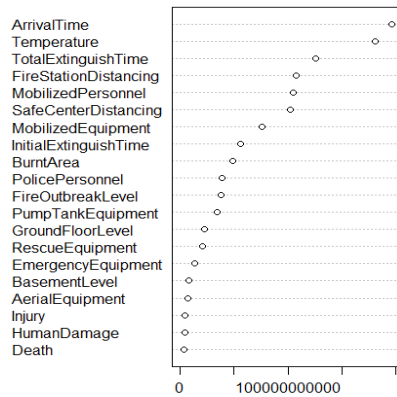


Fig. 6. Importance of independent variables: Detached house fires

Table 5. Analysis result: Detached house fires

Fire factor	%IncMSE	IncNodePurity
Death	9.48E+05	3.2E+09
Injury	4.79E+05	4.49E+09
HumanDamage	1.44E+06	4.23E+09
ArrivalTime	7.90E+06	1.96E+11
InitialExtinguishTime	3.17E+07	5.56E+10
TotalExtinguishTime	5.88E+07	1.26E+11
GroundFloorLevel	1.00E+07	2.2E+10
BasementLevel	-2.34E+06	8.09E+09
FireOutbreakLevel	-2.64E+07	3.79E+10
FireStationDistancing	3.55E+07	1.08E+11
SafeCenterDistancing	3.48E+07	1.02E+11
BurntArea	2.72E+07	4.89E+10
MobilizedPersonnel	4.82E+07	1.05E+11
PolicePersonnel	1.07E+07	3.84E+10
MobilizedEquipment	3.53E+07	7.59E+10
PumpTankEquipment	2.73E+07	3.44E+10
AerialEquipment	4.47E+06	7.34E+09
RescueEquipment	3.16E+07	2.04E+10
EmergencyEquipment	7.19E+06	1.32E+10
Temperature	-1.07E+07	1.81E+11

Table 5는 단독주택 화재에서 재산피해에 대해 화재 인자들의 영향 정도를 나타내는 정확도와 중요도를 확인할 수 있다. 이때 중요도 값이 클수록 재산피해에 미치는 영향이 증가함을 의미한다.

Fig. 6은 Table 5의 분석결과로부터 단독주택 화재에서 재산피해에 대한 독립 변수인 화재 인자의 중요도를 시각화하였다.

Fig. 6에서 단독주택 화재에서 재산피해에 대한 화재 인자 중요도는 출동소요시간, 온도, 화재진압시간, 소방서거리, 동원인력, 안전센터거리, 동원장비, 화재초진시간 등의 순으로 분석되었다.

기타주택(Etc house)에서의 화재로 인한 재산피해에 영향을 미치는 인자들의 영향에 관한 랜덤포레스트 분석 결과는 Table 6과 같이 산출되었다. 이때 예측 모델의 평균제곱근오차는 14,649로 나타났다.

Table 6. Analysis result: Etc house fires

Fire factor	%IncMSE	IncNodePurity
Death	-3.42E+04	1.01E+07
Injury	-4.21E+04	1.11E+07
HumanDamage	-1.82E+04	2.14E+07
ArrivalTime	-2.85E+06	1.02E+10
InitialExtinguishTime	2.92E+07	3.32E+09
TotalExtinguishTime	2.37E+07	5.84E+09
GroundFloorLevel	5.55E+06	5.65E+08
BasementLevel	9.96E+05	4.62E+07
FireOutbreakLevel	1.50E+05	2.46E+08
FireStationDistancing	7.41E+06	6.40E+09
SafeCenterDistancing	9.65E+06	5.09E+09
BurntArea	6.30E+06	5.39E+09
MobilizedPersonnel	1.55E+07	4.95E+09
PolicePersonnel	2.90E+06	2.13E+09
MobilizedEquipment	1.80E+07	2.88E+09
PumpTankEquipment	1.95E+07	2.51E+09
AerialEquipment	2.67E+05	2.46E+08
RescueEquipment	-2.14E+06	1.75E+09
EmergencyEquipment	-4.56E+05	4.10E+08
Temperature	-1.83E+06	7.25E+09

Table 6은 기타주택 화재에서 재산피해에 대해 화재 인자들의 영향 정도를 나타내는 정확도와 중요도를 확인할 수 있다. 이때 중요도 값이 클수록 재산피해에 미치는 영향이 증가함을 의미한다.

Fig. 7은 Table 6의 분석결과로부터 기타주택 화재에서 재산피해에 대한 독립 변수인 화재 인자의 중요도를 시각화하였다.

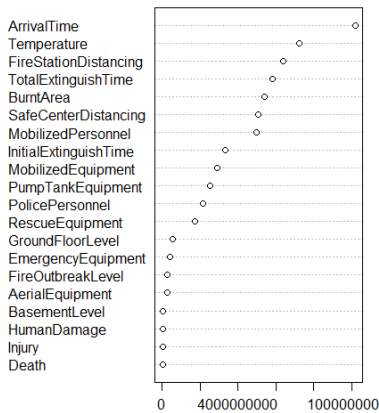


Fig. 7. Importance of independent variables: Etc house fires

Fig. 7에서 기타주택 화재에서 재산피해에 대한 화재 인자 중요도는 출동소요시간, 온도, 소방서거리, 화재진압시간, 소실면적, 안전센터거리, 동원인력, 화재초진시간 등의 순으로 나타났다.

주거지 유형에 따라 재산피해의 화재 영향도 비교를 위해 Table 4-6과 Fig. 5-7에서 제시한 화재 인자별 중요도 수치를 Table 7에 나타내었다. Table 7에서 주거지 유형과 관계없이 출동소요시간, 소방서거리, 화재진압시간은 가장 높은 영향을 미치는 화재 인자임을 알 수 있다. 이어 공동주택의 경우 발화층, 다중주택의 경우 동원인력, 기타주택의 경우 소실면적이 다음으로 높은 중요도를 보였다.

상기와 같은 결과에서 출동소요시간, 소방서거리, 화재진압시간은 모든 주택화재에서 재산피해의 최소화를 위해 공통으로 대응방안이 마련되어야 하는 화재 인자라는 것을 의미한다. 그리고 공동주택 화재에서 중요도를 보인 발화층은 최근 고층화 추세를 보이는 아파트 및 주상복합아파트 등이 공동주택 유형에 포함되기 때문에 파악된다. 단독주택 화재의 경우 동원인력이 중요한 화재 인자로 분석된 것은 타 주택 유형과 비교하면 단독주택의 세부 유형들인 다가구주택, 다중주택, 단독주택, 상가주택 등의 화재에서는 신속한 인력 동원을 통한 초기 화재진압이 요구된다는 것을 알 수 있다. 기타주택 화재의 경우 기타주택의 세부 유형인 컨테이너나 비닐하우스와 같은 주거 형태의 화재에서는 공동주택이나 단독주택 유형과 비교할 경우 소실면적 자체가 재산피해와 직접적인 상관성을 나타낸다. 이로부터 기타주택에서의 소실면적 화재 인자가 공동주택이나 단독주택 유형보다 상대적으로 높은 중요도를 가지는 것으로 분석된다.

Table 7. Importance of independent variables

Independent variables	Importance		
	Multi-unit	Detached	Etc
Death	2.71E+09	3.2E+09	1.01E+07
Injury	2.52E+09	4.49E+09	1.11E+07
HumanDamage	2.83E+09	4.23E+09	2.14E+07
ArrivalTime	1.02E+11	1.96E+11	1.02E+10
InitialExtinguishTime	5.78E+10	5.56E+10	3.32E+09
TotalExtinguishTime	9.6E+10	1.26E+11	5.84E+09
GroundFloorLevel	5.31E+10	2.2E+10	5.65E+08
BasementLevel	1.12E+10	8.09E+09	4.62E+07
FireOutbreakLevel	9.25E+10	3.79E+10	2.46E+08
FireStationDistancing	1.26E+11	1.08E+11	6.40E+09
SafeCenterDistancing	7.14E+10	1.02E+11	5.09E+09
BurntArea	5.35E+10	4.89E+10	5.39E+09
MobilizedPersonnel	6.98E+10	1.05E+11	4.95E+09
PolicePersonnel	6.58E+10	3.84E+10	2.13E+09
MobilizedEquipment	5.02E+10	7.59E+10	2.88E+09
PumpTankEquipment	4.07E+10	3.44E+10	2.51E+09
AerialEquipment	1.37E+10	7.34E+09	2.46E+08
RescueEquipment	1.47E+10	2.04E+10	1.75E+09
EmergencyEquipment	1.43E+10	1.32E+10	4.10E+08
Temperature	8.25E+10	1.81E+11	7.25E+09

이러한 분석결과는 Table 3에서 제시한 피어슨 상관관계 분석결과와 비교하여 화재진압시간, 화재초진시간, 동원인력 등은 주택화재에서 재산피해에 미치는 주요한 화재 인자라는 것을 공통으로 확인할 수 있었다. 다만, 소실면적 화재 인자의 경우 피어슨 상관관계 분석결과에서는 가장 높은 선형적 상관성을 보이는 것으로 나타났다. 반면에 랜덤포레스트 예측 모델에 따른 분석결과에서는 상대적으로 낮은 중요도를 나타내는 결과를 보였다.

Table 7은 주거지 유형별 주택화재에서 화재 인자들이 재산피해에 미치는 중요도를 독립적으로 분석한 결과로, 주거지 유형별로 독립 변수들의 중요도 및 변화 정도가 일정하지 않다. 이에 주거지 유형별로 가장 높은 중요도를 보인 공동주택(Multi-unit house)의 소방서거리(FireStationDistancing), 단독주택(Detached house)과 기타주택(Etc house)의 출동소요시간(ArrivalTime) 독립 변수를 기준으로 나머지 독립 변수들의 상대적 중요도를 백분율로 변환하여 분석한 결과를 Fig. 8에 나타내었다.

Fig. 8에서 소방서거리(FireStationDistancing)는 공동주택(100%)의 중요도가 단독주택(55.1%)이나 기타주택(62.7%)이 비해 높게 나타났다. 출동소요시간(ArrivalTime)은 단독주택(100%)과 기타주택(100%)에 비해 공동주택(81%)이 낮은 결과를 보였다. 이는 소방서거리가 출동소요시간에 비해 낮지 않으므로, 신속한 화재진압을 위해 공동주택 화재 발생시 도착시간을 최소화할 수 있는 교



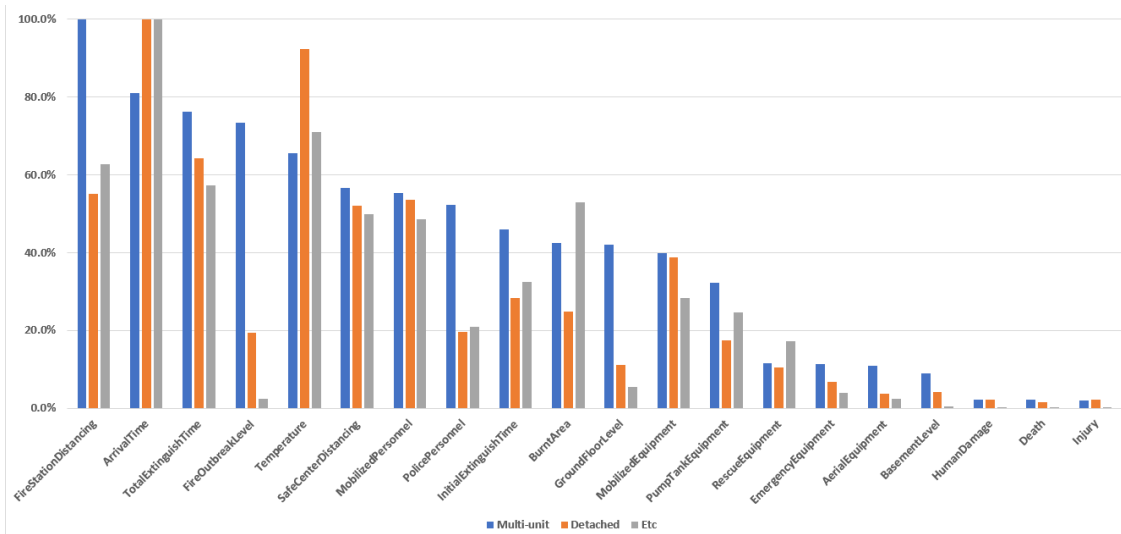


Fig. 8. Relative importance of independent variables by house types

통체계 및 출동지연 해소, 그리고 소화기구 설치 강화 등을 위한 제도적인 노력이 요구됨을 알 수 있다. 화재진압 시간(TotalExtinguishTime)은 공동주택(76.2%), 단독주택(64.3%), 기타주택(57.3%)로 주거지 유형에 따른 큰 차이를 보이지 않는다. 발화층(FireOutbreakLevel)은 단독주택(19.3%)과 기타주택(2.4%)에 비해 공동주택(73.4%)에서 매우 높은 중요도를 나타낸다. 따라서 고층화 추세를 보이는 공동주택의 아파트와 주상복합아파트 등에서 고층 주거지에 대한 화재 예방 대책이 필요함을 알 수 있다. 온도의 경우 단독주택(92.3%)에서의 중요도가 공동주택(65.5%)이나 기타주택(71.1%)보다 높은 경향을 나타내므로 계절적 요인에 대한 추가 분석이 요구된다. 경찰동원명수(PolicePersonnel)는 공동주택(52.2%)에서 상대적으로 높은 중요도를 보이는데, 이는 화재 피해를 줄이기 위한 직접적인 영향이라기보다는 화재 규모에 따라 화재진압 인력과 장비의 현장 투입을 지원하는 간접적인 영향을 미치는 것으로 분석된다. 화재초진시간(InitialExtinguishTime)의 경우에는 공동주택(45.9%)에서 단독주택(28.4%)이나 기타주택(32.5%)보다 높은 중요도를 나타낸다. 고층 주거지가 많은 공동주택에서의 화재초진시간이 재산피해에 미치는 영향이 크다는 것을 의미하므로 공동주택에서의 화재초진시간을 줄이기 위한 화재 예방 활동이 필요함을 알 수 있다. 소실면적(BurntArea)의 경우 화재로 인해 기타주택(52.8%), 공동주택(42.5%), 단독주택(24.9%)의 순서로 소실면적에 따른 재산피해가 크다는 것을 의미하므로 화

재초진시간 및 화재진압시간을 감소시키기 위한 노력과 연계하여 대책 마련이 요구된다. 건물층수\_지상(GroundFloorLevel)은 주거지 유형별로 발화층과 유사한 수준의 중요도를 보이므로, 발화층에서 기술한 바와 같이 고층 주거지에 대한 화재 예방 및 대책 강화가 필요하다. 이어 동원장비소계(MobilizedEquipment), 펌프물탱크(PumpTankEquipment), 구조구급장비(RescueEquipment), 구급(EmergencyEquipment), 고가\_굴절(AerialEquipment), 건물층수\_지하(BasementLevel), 인명피해(HumanDamage), 사망(Death), 부상(Injury)의 경우 주거지 유형에 따른 변별력은 크지 않은 것으로 나타났다.

#### 4. 결론

제안 연구는 주택화재에서 공동주택, 단독주택, 기타주택 등 주거지 유형에 따른 재산피해에 영향을 미치는 주요 화재인자를 평가하고 이에 관한 제도적 개선을 위한 빅데이터 기반의 과학적 근거를 제시하였다. 데이터 분석을 위해 소방청에서 제공하는 2017년도 화재 데이터를 이용하였다. 화재인자 영향평가 결과에서 공동주택에서는 소방서거리, 출동소요시간, 화재진압시간, 발화층의 순으로 중요도가 나타났다. 단독주택과 기타주택은 출동소요시간, 온도, 화재진압시간, 소방서거리 등의 순으로 높은 중요도를 보였다. 이로부터 재산피해를 최소화

하기 위해서는 조기에 화재진압이 필요함을 알 수 있었다. 화재진압시간을 줄이기 위해서는 소방서거리가 가깝고 출동소요시간은 작을수록 유리하다. 이에 화재진압을 위한 인력과 장비가 조기에 현장에 접근할 수 있도록 교통 통제나 현장 접근을 지연하는 각종 요인을 법과 제도적으로 보완할 필요가 있다는 결론을 얻었다. 또한, 공동주택의 경우 최근 고층아파트나 주상복합아파트의 화재로 인해 화재 발생층이 중요한 인자로 분석되었다. 그러므로 고층에서 발생하는 화재를 조기에 진압하기 위한 자동소화설비 설치 및 운영 강화를 위한 조치가 필요함을 알 수 있었다.

향후 연구로는 중요 요인으로 도출된 화재 인자들에 관한 개선 방안을 고도화하기 위해 해당 주택들의 행정 및 교통 정보, 소방시설 등에 대한 영향을 고려할 필요가 있다. 그리고 단독주택과 기타주택에서 높은 중요도를 보인 온도 화재 인자와 관련하여 화재의 계절적 요인에 대한 추가적인 데이터 확보 및 분석이 요구된다.

## References

- [1] B. Kim, C. Seo, W. G. Shin, "Numerical study of the factors affecting fire flow velocity in the case of interior fire in an apartment building", *Journal of Fire Science Engineering*, Vol. 30, No. 4, pp. 14-19, Aug. 2016.  
DOI: <https://doi.org/10.7731/KIFSE.2016.30.4.014>
- [2] J. Kim, H. Kong, "Multivariate analysis of fire prevention activities, special investigations of fire safety, and fire safety management by the apartment buildings management methods and inspection of firefighting facilities" *Journal of the Convergence on Culture Technology*, Vol. 6, No. 2, pp. 489-502, May 2020.  
DOI: <https://doi.org/10.17703/JCCT.2020.6.2.489>
- [3] National Fire Agency, 2020 National Fire Agency Fire Statistical Year Book, July 2020.
- [4] J. Ku, "A study on evaluation plan of fire safety performance for public building", *Journal of the Korea Convergence Society*, Vol. 10, No. 4, pp. 25-32, Sept. 2019.  
DOI: <https://doi.org/10.15207/JKCS.2019.10.4.025>
- [5] S. Jeong, Y. Park, J. Kim, "A checklist and manual developed to review fire safety facilities' compliance with fire safety requirements for apartment buildings", *Korean Journal of Construction Engineering and Management*, Vol. 19, No. 6, pp. 94-102, Nov. 2018.  
DOI: <https://doi.org/10.6106/KJCEM.2018.19.6.094>
- [6] E. Hwang, D. Seo, "A study on identifying problems by the comparative analysis of high-rise apartment evacuation design codes", *Journal of Fire Science and Engineering*, Vol. 31, No. 6, pp. 47-52, Nov. 2017.  
DOI: <http://dx.doi.org/10.7731/KIFSE.2017.31.6.047>
- [7] I. Lee, *Risk prediction and analysis of building fires-based on property damage and occurrence of fires-*, Master's thesis, Yonsei University Business Analytics, Seoul, Korea, 2020.
- [8] W. Jeong, "A study on the impact assessment of fire factors influencing the damage of human life in residential fires", *Journal of Knowledge Information Technology and Systems*, Vol. 16, No. 3, pp. 471-481, June 2021.  
DOI: <http://doi.org/10.34163/ikits.2021.16.3.009>
- [9] W. Jeong, "A study on fire factors analysis influencing property damage by the use of fire extinguishing equipment in building and structure fires", *Journal of the Korea Academia-Industrial cooperation Society*, Vol. 22, No. 8, pp. 138-145, Aug. 2021.  
DOI: <http://doi.org/10.5762/KAIS.2021.22.8.138>
- [10] Open Data Portal, National Fire Agency\_Address information of outbreak of fire, <https://www.data.go.kr/data/15044005/fileData.do>, Oct. 2020.
- [11] R. L. Wasserstein, A. L. Schirm, N. A. Lazar, "Moving to a world beyond "p < 0.05"", *The American Statistician*, Vol.73, pp.1-19, 2019.  
DOI: <http://dx.doi.org/10.1080/00031305.2019.1583913>
- [12] Y. Park, S. Jeon, T. Y. Kwon, "A sample size calibration approach for the p-value problem in huge samples", *Communications for statistical applications and methods*, Vol.25, No.5, pp.545-557, 2018.  
DOI: <https://doi.org/10.29220/CSAM.2018.25.5.545>
- [13] K. Pearson, *Mathematical contributions to the theory of evolution. III. Regression, heredity and panmixia*, *Philosophical transactions of the Royal Society A*, Vol.187, pp.253-318, 1896.
- [14] A. J. Bishara, J. B. Hittner, "Testing the significance of a correlation with nonnormal data: Comparison of Pearson, Spearman, transformation, and resampling approaches", *Psychological methods*, Vol. 17, No. 3, pp. 399-417, 2012.  
DOI: <http://dx.doi.org/10.1037/a0028087>
- [15] L. Breiman, "Random forests", *Machine Learning*, Vol. 45, pp. 5-32, 2001.
- [16] V. Svetnik, A. Liaw, C. Tong, T. Wang, "Application of Breiman's random forest to modeling structure-activity relationships of pharmaceutical molecules", *Lecture Notes in Computer Science*, Vol. 3077, pp. 334-343, 2004.  
DOI: [http://dx.doi.org/10.1007/978-3-540-25966-4\\_33](http://dx.doi.org/10.1007/978-3-540-25966-4_33)
- [17] K. S. Ko, D. H. Hwang, S. J. Park, G. G. Moon, "Electrical fire prediction model study using machine learning", *Journal of Korea Institute of Information*,

Electronics, and Communication Technology, Vol. 11,  
No. 6, pp. 703-710, 2018.  
DOI: <http://dx.doi.org/10.17661/jkiect.2018.11.6.703>

---

정 원 일(Weonil Jeong)

[정회원]



- 1998년 2월 : 인하대학교 전자계산공학과 (공학사)
- 2004년 8월 : 인하대학교 컴퓨터정보공학과 (공학박사)
- 2004년 7월 ~ 2006년 7월 : 한국전자통신연구원 선임연구원
- 2007년 3월 ~ 현재 : 호서대학교 컴퓨터공학부 교수

〈관심분야〉

데이터베이스, 빅데이터분석, 시스템 보안