
빅데이터 기반 전기화재 예측 모델을 활용한 데이터 유통 서비스 개발

- 선도소프트 컨소시엄(한국전기안전공사, DB손해보험) -

2018. 11. 20



Contents

I 사업개요

II 사업목표

III 빅데이터 구축 현황

IV 사업추진현황

V 분석 방법론

VI 분석 결과

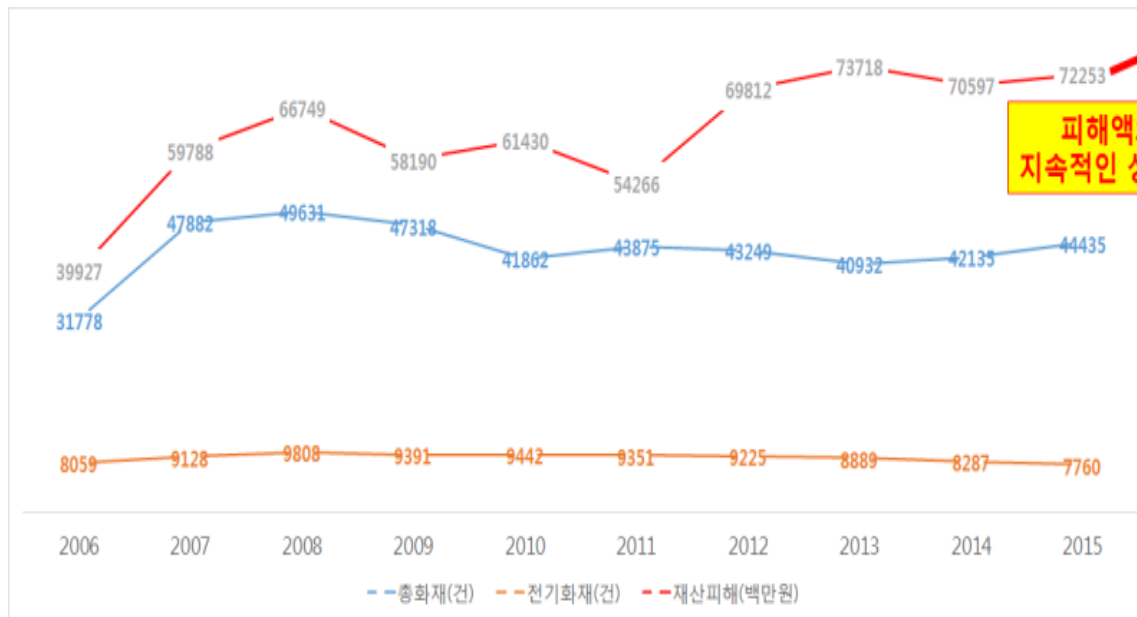
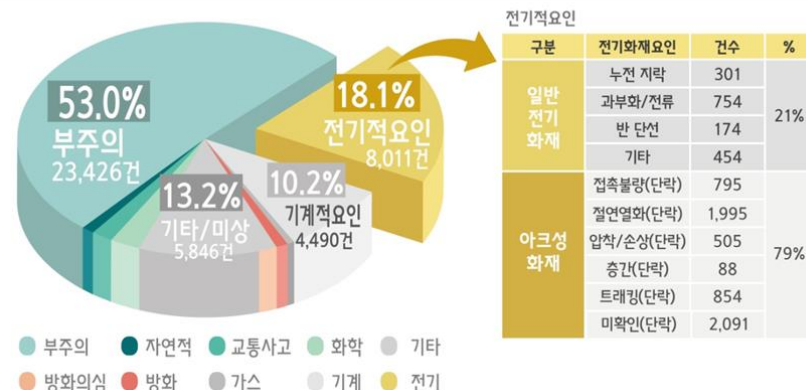
VII 검증

VIII 기대효과

추진 배경

피해액의 규모가 날로 커지는 **전기 화재**를
빅데이터로 분석하고 이를 국민들에게
 체감할 수 있고, 불의의 사고에 대비한 **건물화재보험상품**
 적용 필요

2017 화재 요인 별 분석



- ❖ 과제명
- ❖ 수행기관
- ❖ 사업기간
- ❖ 수행내용

- ❖ 최종목표

빅데이터 기반 전기화재 예측 모델을 활용한 데이터 유통 서비스 개발

(주관기관) (주)선도소프트 / (참여기관) 한국전기안전공사, DB손해보험 주식회사

2018. 6. 28 ~ 2018. 12. 20

- 빅데이터 기반 전기화재 예측 모델을 활용한 데이터 유통서비스

- 1) 전기 화재 및 안전점검 DB, 건축물 데이터 수집/융합 및 분석(전국규모)
- 2) 빅데이터 기반 인공지능 예측 모델 및 전기안전등급제 고도화(화재발생위험도)
- 3) 손해보험 요율산정 고객 맞춤형 화재보험 상품 개발

* 일반용 전기 : 전압 660V 이하로 용량 75kW 미만 전기설비

데이터 기반 행정 활성화 및 국가 중점 데이터 개방 정책과
국민 안전에 대한 공적 역할을 고려한 **실질적인 국민체감 서비스**
전기화재 위험 지수 도출과 이에 따른 맞춤형 화재보험 상품개발



국민의 전기화재안전 지킴이



전기안전 빅데이터 제공 및 검증

데이터 제공 및 현장검증



- ▶ 전기재해(화재·감전·설비)에 대한 **사고현장 조사** ('01년 ~ 계속, 전력기반센터)
- ▶ **일반용 전기설비 안전점검** ('01년 ~ 계속, 전력기반센터)
- ▶ **사물인터넷 기반 국가전기안전관리 기술 개발** (2016년, 전력기반센터)



알고리즘 고도화 및 결과물 시각화

빅데이터 수집·분석 /
인공지능 알고리즘 고도화 /
가시화 시스템 개발



- ▶ 전문 빅데이터 기술력 보유
 - 빅데이터 기반 전기재해 예방예측 알고리즘을 적용한 사고 예방 시스템
 - 대용량 사물인터넷(IoT) 데이터 수집 및 실시간 모니터링 지리정보 (GIS)시스템
 - 실시간 빅데이터 기반 수집, 분석 및 가시화 시스템



민간활용 서비스 구축·제공

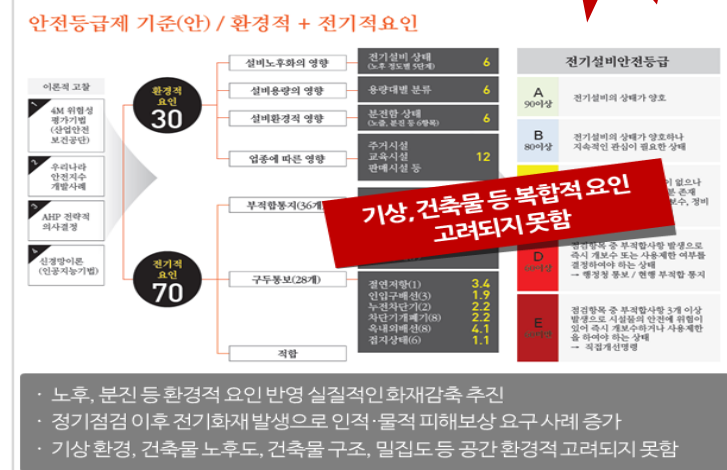
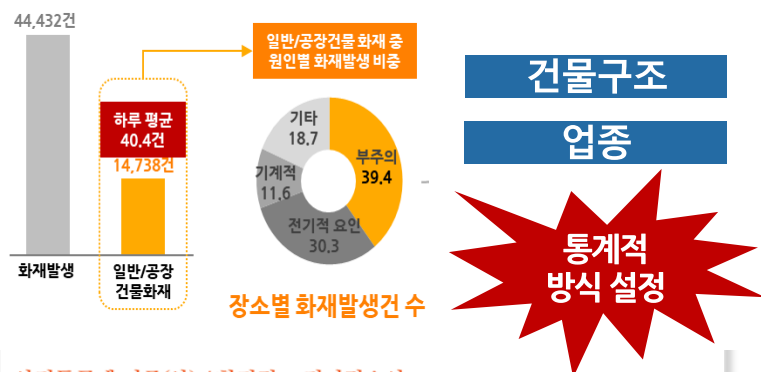
손해요율 적용 및 화재보험 상품 개발



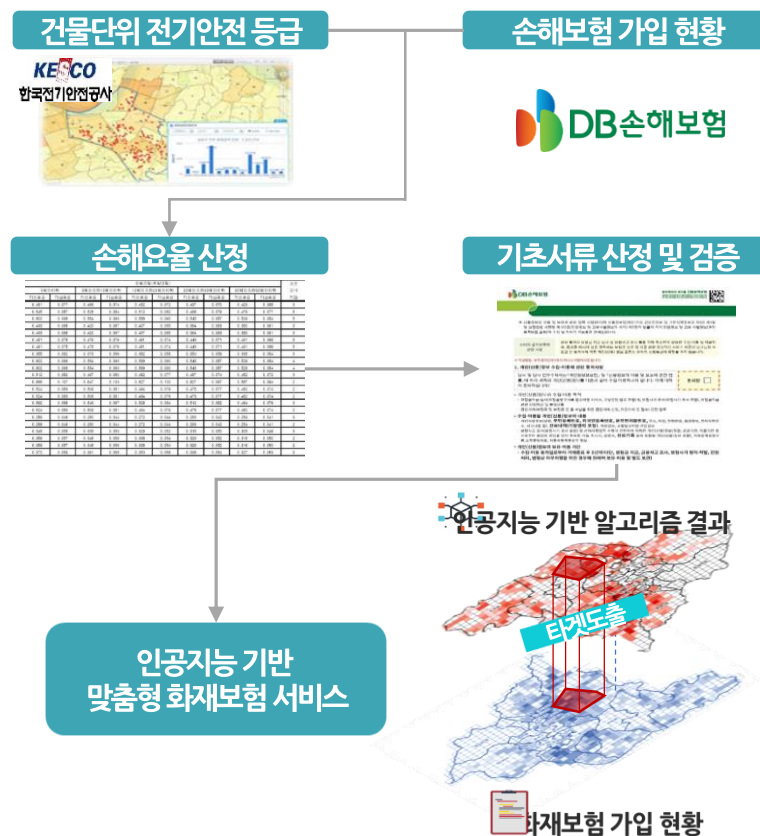
- ▶ 업계최초 SK텔레콤 제휴를 통한 '스마트(SmarT)-UBI 안전운전 특약 자동차 보험'상품 출시
- ▶ T맵과 연계한 안전운전 점수 대비 가입자 할인 제공

AS-IS / TO-BE : 손해보험 요율산정 고객 맞춤형 화재보험 상품 개발

AS-IS) 일률적인 보험료 산정



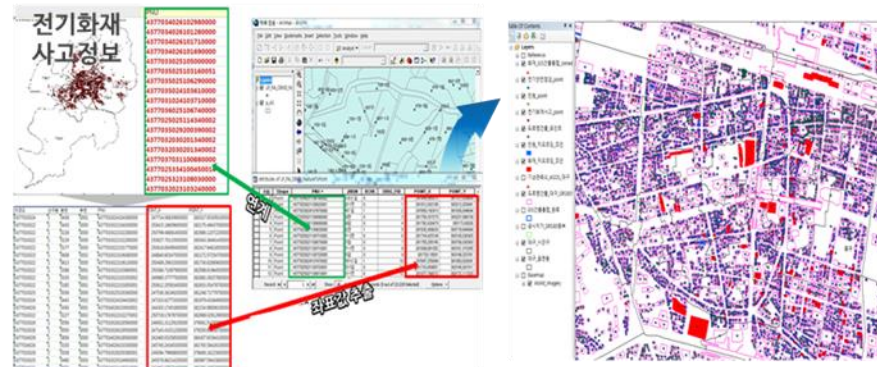
TO-BE) 빅데이터 기반 화재보험 상품개발



최초 전국 규모 전기화재 사고와 공동 데이터 융합(1)

Comments	Histogram	NumDistinct	NumNulls	Comments	Histogram	NumDistinct	NumNulls
자료구분	Frequency	3	0	오염물	None	38	2,152
일반범위	None	91,976	0	오염물질상속	None	51	2,459
관리자정보여부	None	2	0	오염물질	None	1,104	2,574
등록일자	None	2,715	0	대상도로명	Height Balanced	64,508	5,040
등록자	None	1	0	오염도	None	475	5,312
조사서번호	Height Balanced	93,045	17	올로드 시	None	17	6,344
소방서명	Frequency	194	17	None	None	218	6,344
화재발생일시	Height Balanced	93,040	17	발화원인기대분류	Frequency	159	9,852
시간대	Frequency	12	17	가을 지화물	None	17	12,034
요일	Frequency	7	17	올로드 구	None	103	15,719
상세주소	None	76,360	17	보통공역 현황	None	915	16,226
화재유형	Frequency	4	17	올로드 읍면동	None	119	16,704
발화원인대분류	None	8	17	발화원인기대분류	None	15	17,354
발화원인소분류	None	20	17	발화원인대분류	Frequency	2	28,915
발화원인소분류	Frequency	10	17	시각 범위	None	2,553	30,928
최초발화대분류	None	13	17	올로드 리	None	38	37,698
최초발화소분류	None	74	17	주변범위	None	21,042	38,100
동작원인대분류	None	6	17	물 범위	None	641	45,586
동작원인소분류	None	24	17	수전용량	None	739	50,333
연소확대대분류	None	13	17	안전관리주제	Frequency	12	51,242
연소확대소분류	None	79	17	안전전안	Frequency	14	51,287
연소확대사유	None	12	17	화재발생사실과	None	4	52,093
인명피해소계	None	17	17	최종결정사실	None	3,496	52,238
사망	None	9	17	결정결정사실	None	7	52,958
부상	None	16	17	속보구분	Frequency	5	53,047
재산피해액	None	9,709	17	유일자	Frequency	2	54,652
재산피해소계	None	19,408	17	유일시간 일종차용	None	36,708	55,074
화재보험회수가입대상	None	2	17	유일결정사실	None	5	57,025
별리	None	9	17	None	None	30,498	61,993
건물구분 식	None	8	17	시도	Frequency	18	62,369
건물구분 프	None	2	17	119접화	None	847	64,832
건물구분 중	None	15	17	문속	None	9	70,006
건물상위	None	8	17	리	None	4,632	73,718
장소대분류	Frequency	14	17	발화개요	None	18,828	74,206
장소중분류	Frequency	42	17	화재진입시간	None	1,719	74,546
장소소분류	None	274	17	개조국가명	None	19	76,332
특정소방대일	None	25	17	구분2	None	33	80,968
발화원리대일	None	2	17	최초결정사실	None	440	82,604
다중이용일여부	None	2	17	소유자관계자	None	184	83,754
부속물대분류	None	6	17	유류자관계자	None	145	83,789
부속물소분류	None	36	17	소방서 일련번호	None	4,078	83,917
발화지점대분류	None	8	17	고려범위	None	5,766	86,657
발화지점소분류	None	77	17	개조회사명	None	1,045	89,108
연소확대범위	None	6	17	개조명	None	1,523	89,506
원안공역구분	None	10	17	None	None	2,746	89,844
사상시행인	None	11	17	119지역	None	550	90,880
사상전상위	None	15	17	개조일자	None	795	91,057
누락정보기	None	3	17	사상시행지명	None	658	91,988
시도명	Frequency	17	126	기상정보	None	18	92,461
구분	Frequency	212	137	건축위원회	None	3	93,026
사실소분류	Frequency	64	333	유일일자	None	0	93,045
문속	None	491	425	안전용량구분	None	0	93,045
부속상	None	7,047	749	안전용량구간	None	0	93,045
문속	None	5,820	749	프사자	None	0	93,045
연리동	Height Balanced	3,726	923	전체문로드	None	0	93,045

전처리(오류수정 등)



PK : 건물 고유 번호

국토부

GIS 건축물 통합정보

국토부

건물통합마스터

행안부

도로명주소 건물 데이터

전기안전공사

안전점검일마감
(ES_DAY 등)

97개 필드정보

고객마스터

54개 필드정보

IoT데이터(16개DB)

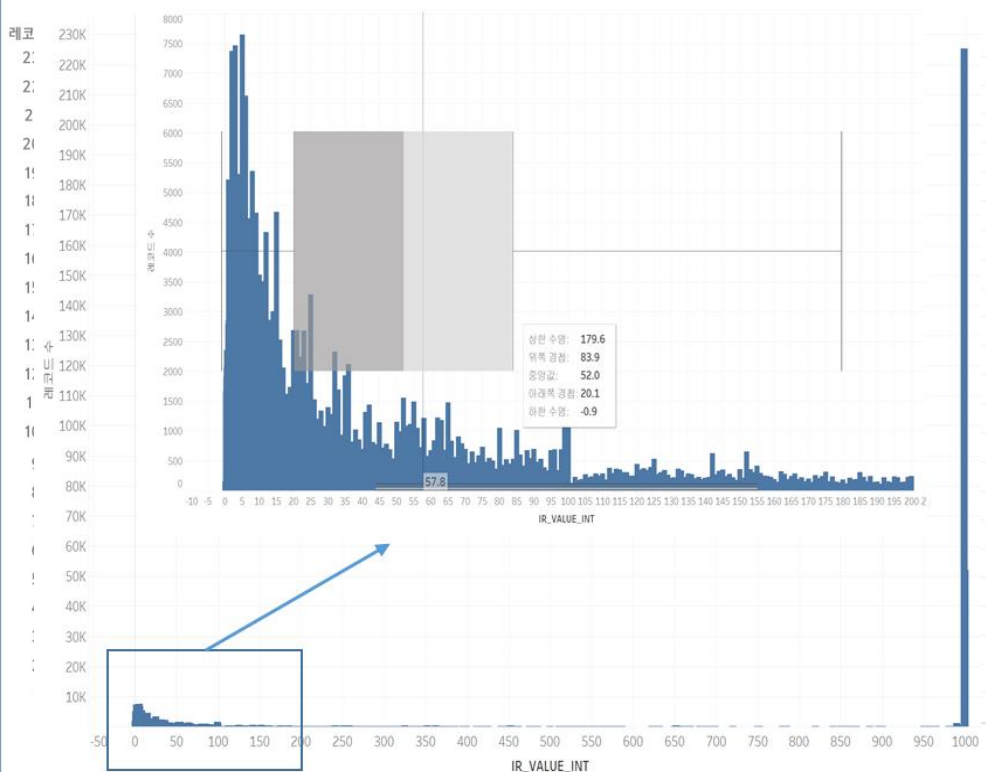
200개 필드정보

전기화재+건물 속성정보+전기점검 정보 융합 데이터를 통한 모델 설계 가능

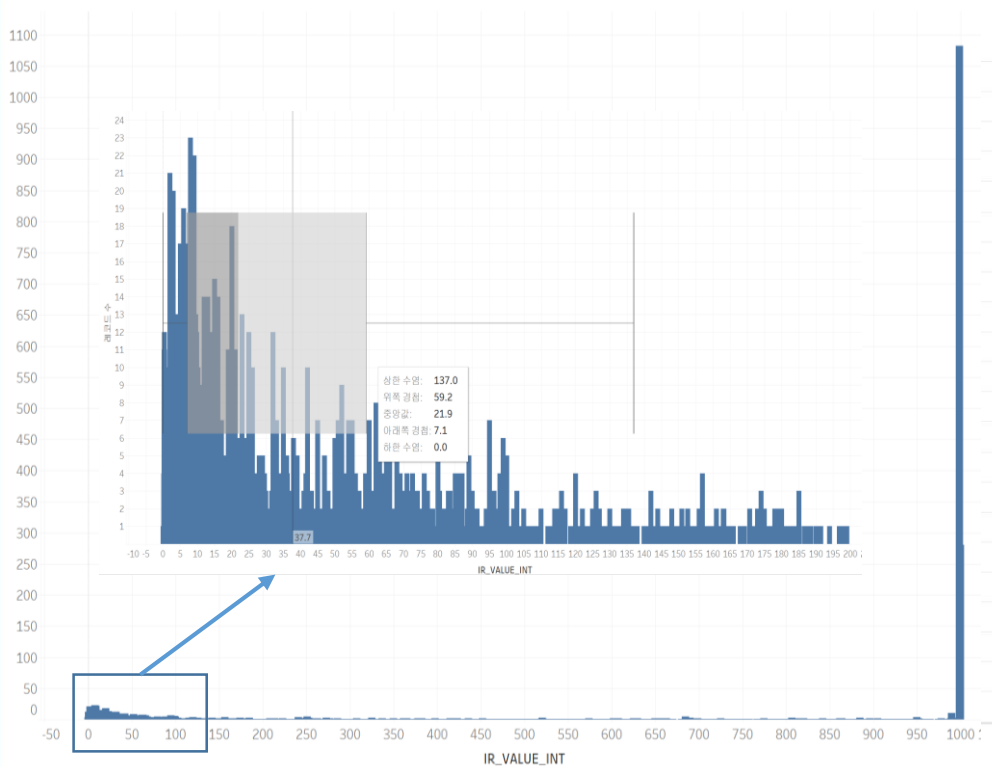
주요 요인 도출 : 절연저항(IR)

아크, 단락 가능성 등 전기화재 발생 요인인 절연 성능을 나타내는 값으로 안전점검 관련 DB에서 가장 잘 관리되고 있으며, 낮을수록 전기화재 발생 가능성이 높다고 알려진 인자. 실제 전기화재 사고/미사고의 차이가 매우 높아 매우 중요한 구분자역할

무사고 IR 분포



사고 IR 분포



주요 요인 도출 : 건물 노후도

건축물의 준공년도를 2017년 12월을 기준으로 역산하여 건물 노후도 파생변수 생성.

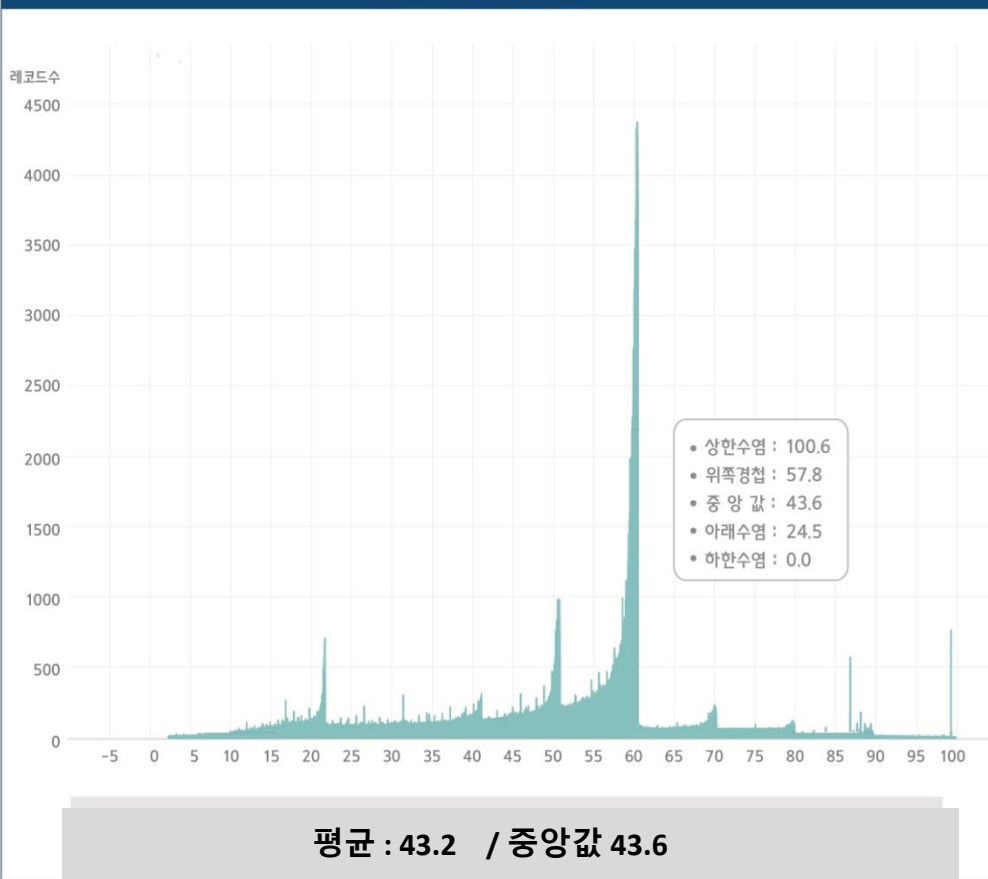
전기화재 건수는 25년 전후가 급격히 떨어지나, 실제로 건물의 수를 대입한 비율을 보면 25년에서 35년 건물에서 높은 비율



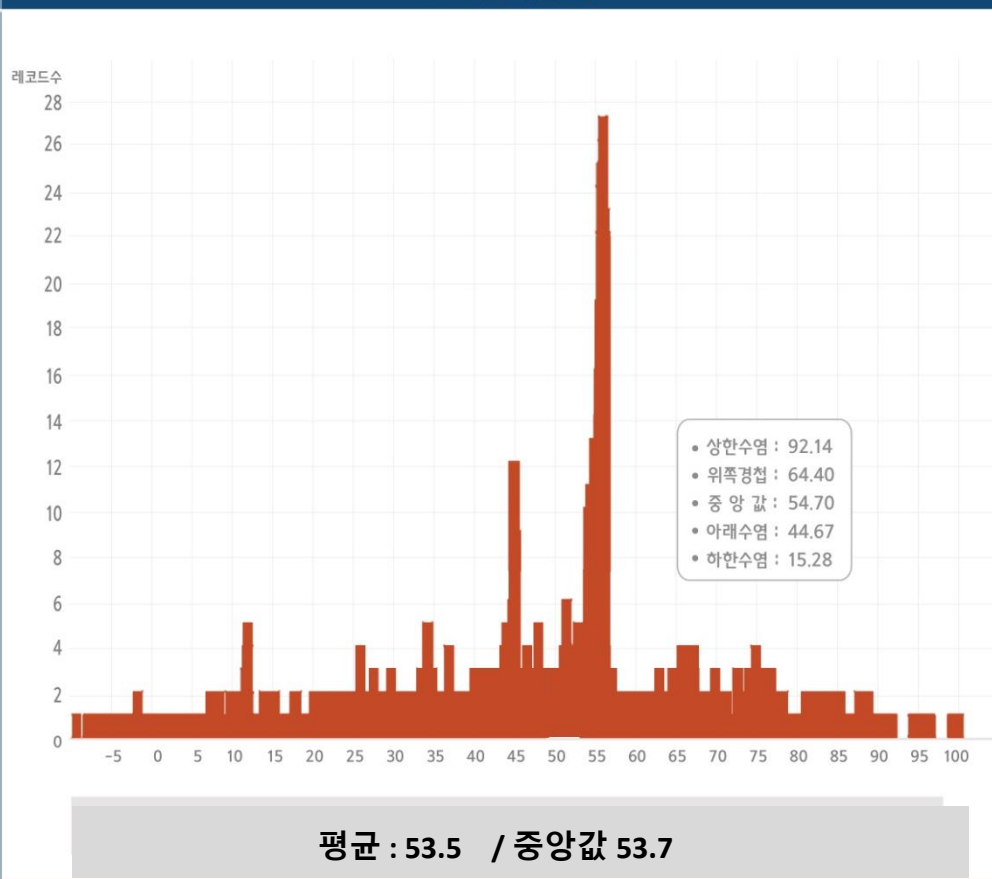
주요 요인 도출 : 건물 노후도

건물 간 밀집도의 지표 중 하나로 건폐율이 높으면 건물 간 밀집도가 높다는 의미로 실제 화재는 건폐율이 높게 나오며 구분자 역할

무사고 건폐율 분포



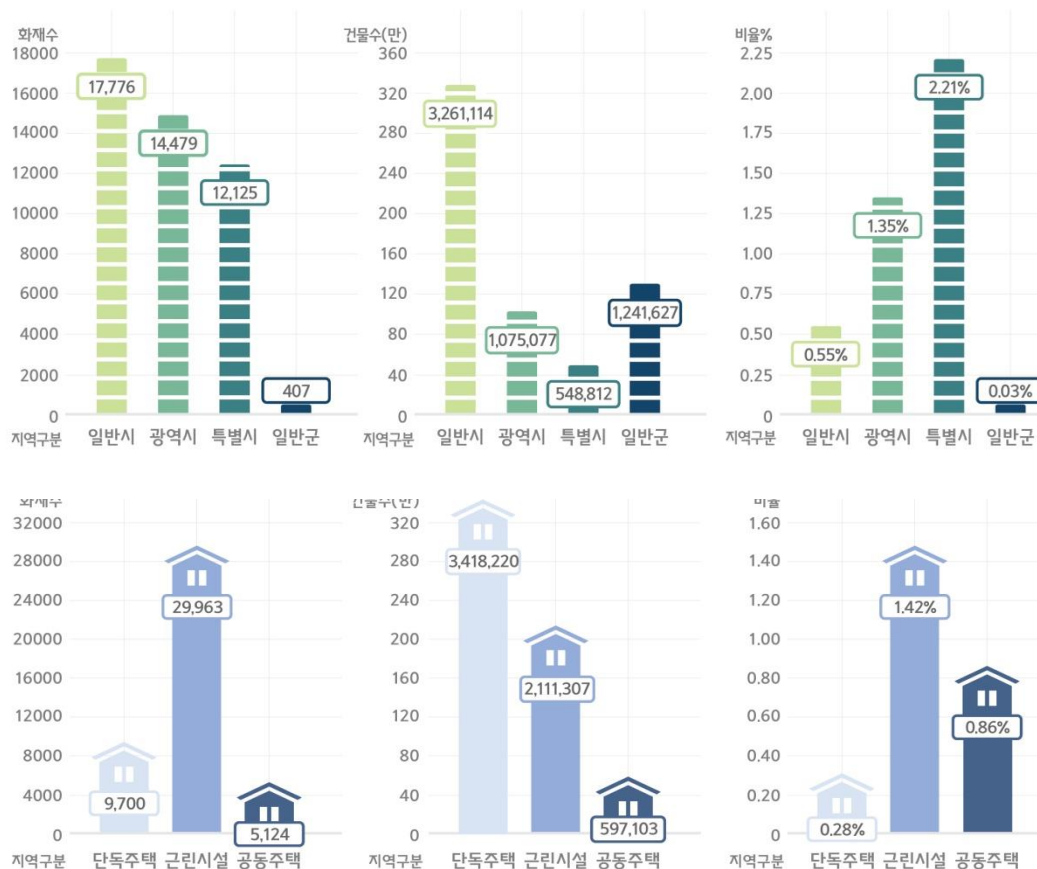
사고 건폐율 분포



* 건폐율 = (건축면적 / 대지면적) × 100 (%)로 지방자치단체의 조례로 정하며, 과밀화, 채광, 화재 대비 등의 목적으로 상이하게 관리

전국규모의 효율적 분석 방법론 제시(1)

지역별 건물화재 특별시-광역시-일반시-일반군으로 높은 비율차가 있어 이를 반영한 알고리즘이 필요.
또한 건물용도별 전기화재 건수도 실제 비율차가 높아 이에 대한 편차 해소 필요



세분화된 최대 12개 알고리즘

	단독주택	공동주택	근린시설
서울시	Ex) 서울시		
광역시	Ex) 대전광역시		
시	Ex) 김해시		
군	Ex) 음성군		

지역 분류 기준

서울특별시

광역시급 : 6개의 광역시(부산, 대구, 인천, 광주, 대전, 울산)

시급 : 75개 자치시

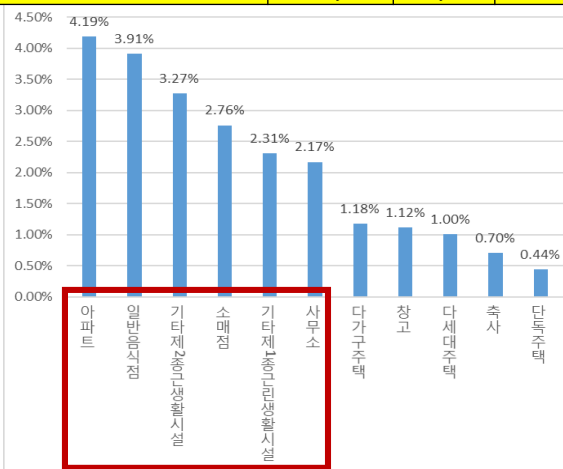
군급 : 77개 군(광역시자체 내 군 제외)

Ⅰ 전국규모의 효율적 분석 방법론 제시(2)

(검토) 실제 지역별로 전기화재는 건물용도 및 건수가 큰 차이를 보이므로 단일모델으로는 한계가 있음

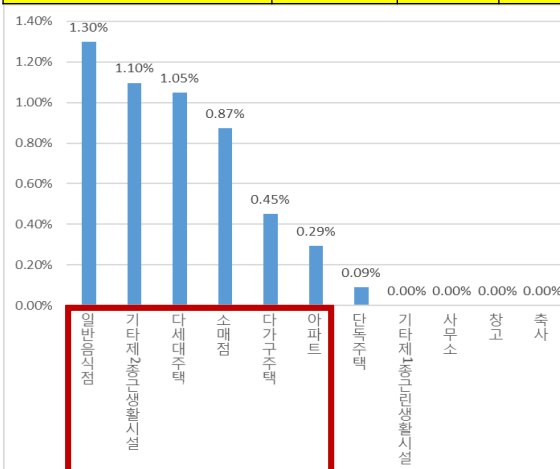
건물당 화재건수 비율 1.09% (대전광역시)

구분	건물수	화재수	화재율
제1종 근린 생활시설	3,377	78	2.3%
제2종 근린 생활시설	2,933	96	3.3%
다가구주택	14,282	168	1.2%
다세대주택	4,384	44	1.0%
단독주택	68,069	298	0.4%
사무소	2,907	63	2.2%
소매점	4,859	134	2.8%
아파트	3,963	166	4.2%
일반음식점	3,426	134	3.9%
창고	1,789	20	1.1%
축사	285	2	0.7%
합계	110,370	1,206	1.1%



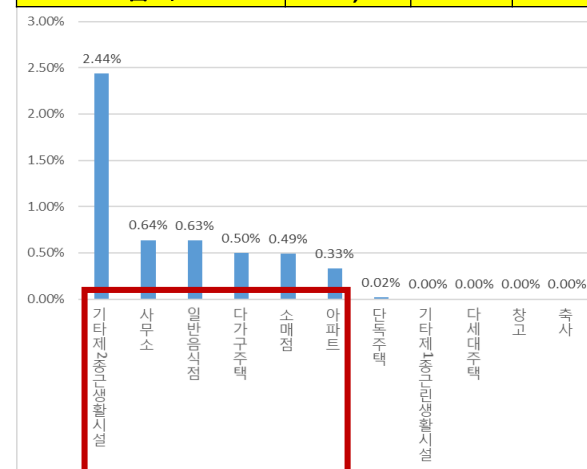
건물당 화재건수 비율 0.27% (김해시)

구분	건물수	화재수	화재율
제1종 근린 생활시설	29		0.0%
제2종 근린 생활시설	1,915	21	1.1%
다가구주택	664	3	0.5%
다세대주택	1,623	17	1.0%
단독주택	33,400	30	0.1%
사무소	95		0.0%
소매점	1,147	10	0.9%
아파트	1,356	4	0.3%
일반음식점	2,848	37	1.3%
창고	237		0.0%
축사	1,726		0.0%
합계	45,040	122	0.3%



건물당 화재건수 비율 0.05% (음성군)

구분	건물수	화재수	화재율
제1종 근린 생활시설	11		0.0%
제2종 근린 생활시설	82	2	2.4%
다가구주택	201	1	0.5%
다세대주택	140		0.0%
단독주택	27,032	5	0.0%
사무소	157	1	0.6%
소매점	204	1	0.5%
아파트	304	1	0.3%
일반음식점	946	6	0.6%
창고	822		0.0%
축사	1,506		0.0%
합계	31,405	17	0.1%

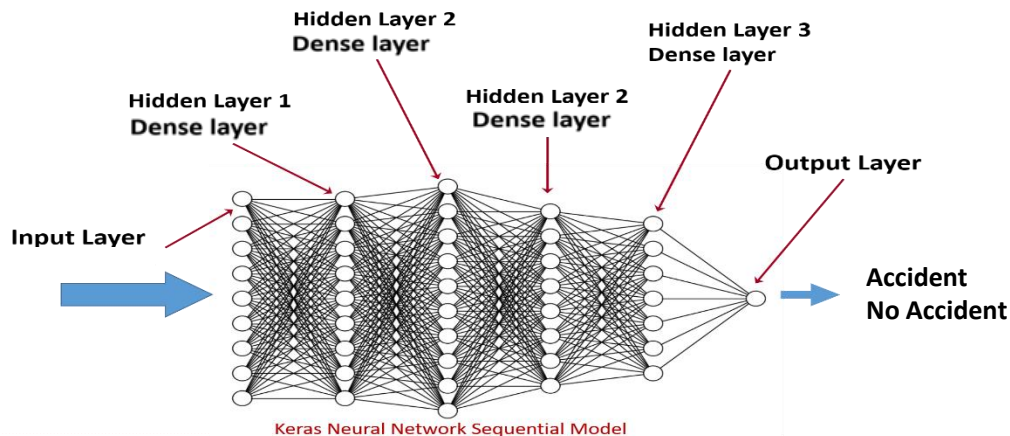


I 인공지능을 활용한 사고 예측 : 정확도 83.3% 달성

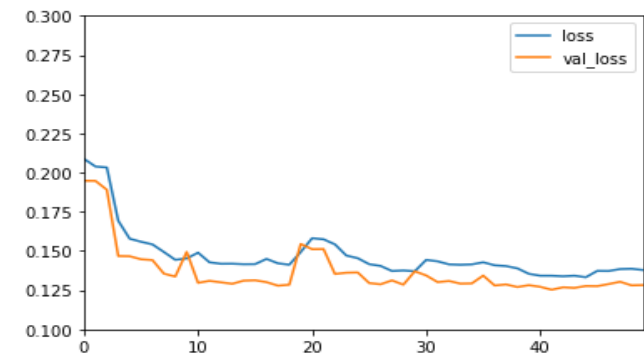
- 14개 설명변수와 1,606 파라미터 값, 12Layer 신경망(Deep Learning) 모델 구축 학습
- TensorFlow 활용 사고/미사고 학습(4만개 학습데이터)하여 **분류 정확도 83.3%(Test Data 적용)**

지역과 용도 별 예측 정확도(ACC = 평균 84.56)

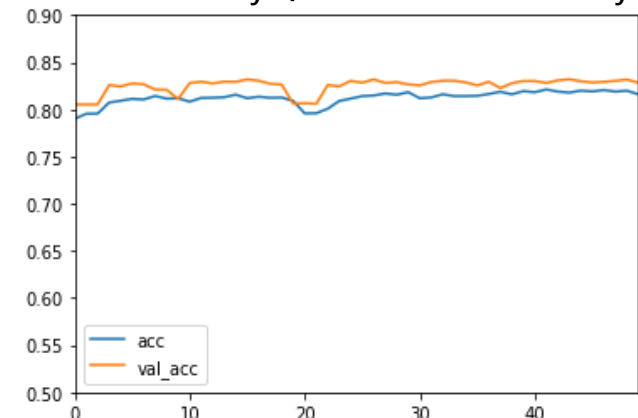
구분	단독주택	공동주택	근린시설
특별시	81.9	91.1	78.7
광역시	80.2	88.9	80.6
일반시	76.3	90.2	84.0
일반군	-	93.7	-



epoch 에 따른 train data loss 와 validation data loss 비교



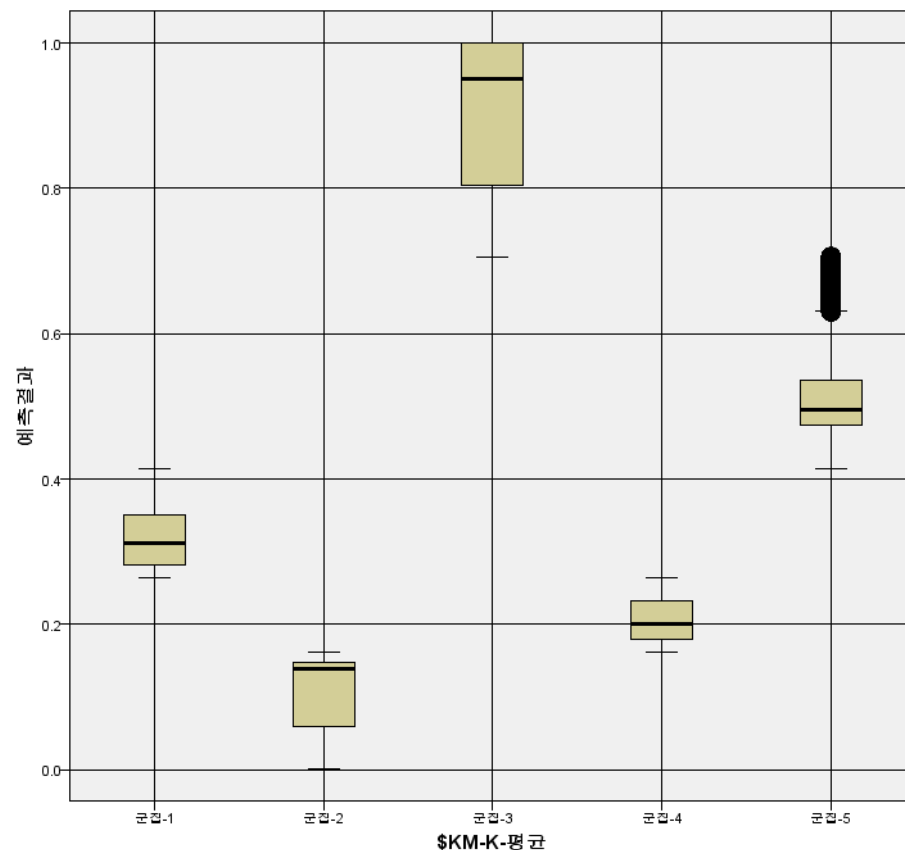
train data accuracy 와 validation data accuracy 비교



전기화재 위험지수(5단계) 도출

- K-means Clustering을 통한 각 건물별 전기화재 위험지수를 5단계(안전,관심,주의,경계,위험) 도출

모델 등급	군집명	집단수	비율
A등급(안전)	군집-2	1,550,794	60.3%
B등급(관심)	군집-4	707,204	27.5%
C등급(주의)	군집-1	84,079	7.2%
D등급(경계)	군집-5	117,585	4.6%
E등급(위험)	군집-3	12,017	0.5%
합계	2,571,679	2,571,679	100.0%



전기화재 위험지수(5단계) 도출

- 인공지능을 통한 건물 지수(5개 등급)에 해당하는 건물 중 실제 3년 간 화재 비율을 도출하면, 12,017개의 건물 중 실제 화재가 913건 발생하여 등급 간 화재 비율은 **위험한 등급에서 72% 전기화재 발생**

모델등급	예측건물수	비율	실제 화재수
A등급(안전)	1,550,794	60.3%	1,604
B등급(관심)	707,204	27.5%	2,492
C등급(주의)	184,079	7.2%	1,417
D등급(경계)	117,585	4.6%	2,011
E등급(위험)	12,017	0.5%	913
합 계	2,571,679	100.0%	8,437

* 실제 화재수 : 15년~17년간 8,437개 적용(출처 : 소방방재청)

선제적 전기재해 예방과 화재보험의 합리화를 통해 대국민 재해 예방 및 피해 비용 절감

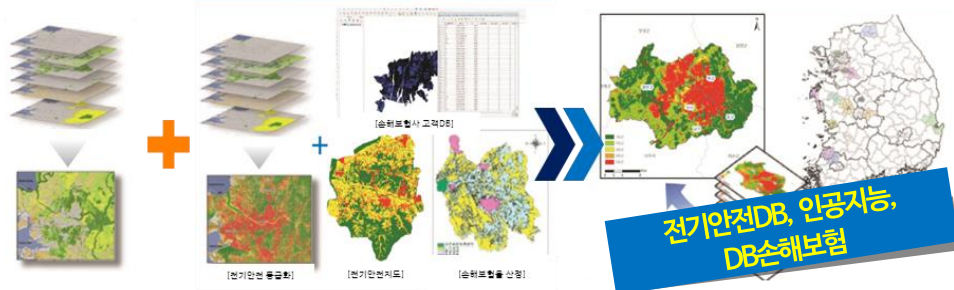
융·복합 新 비즈니스 적용

[손해보험사 연계] 건축물 유형(단독, 아파트, 연립 등)에 따른 전기안전 지수화와 이를 활용한 화재보험료[위험 중심의 보험상품 개발] 산정 실증

전기안전점검DB

타 부처 공공DB + 손해보험DB

전기안전 손해보험율 / 가입자



(전기안전화재지수) 인공지능 기반 예측 알고리즘

(사업화를 위한 전국 확대) 건축 유형별 예측 알고리즘 확대 및 시스템 구축

(건물화재보험 상품개발) 고객 맞춤형 건물화재보험 상품개발 및 확산

전기안전 데이터 활용모델

인공지능
전기안전
등급제

대국민
재해피해
보장확대

손해보험사
직접 참여
상품개발

전기화재
관리 및 보험상품
해외진출 교두보

재해 피해 감소 효과

예산절감

6억 원

12억 원

18억 원

2019년
(개발종료 후 1년)

2020년
(개발종료 후 2년)

2021년
(개발종료 후 3년)

*산출근거: '16년전기안전공사 긴급출동 31.2억원의 20% 절감시

감사합니다.



과학기술정보통신부

NIA 한국정보화진흥원