

# 전세 매물 이상탐지 모델을 통한 전세사기 유형 분석 및 예방 방안 제시

2분반 4조

PART 01

## 서론

1. 주제 선정 배경
2. 분석 방향 및 목표

PART 02

## 모델링 과정

1. 데이터 수집
2. 모델링 설명

PART 03

## 연구 결과

1. 웹사이트 구축

PART 04

## 결론

1. 서비스 활용방안
2. 기대효과

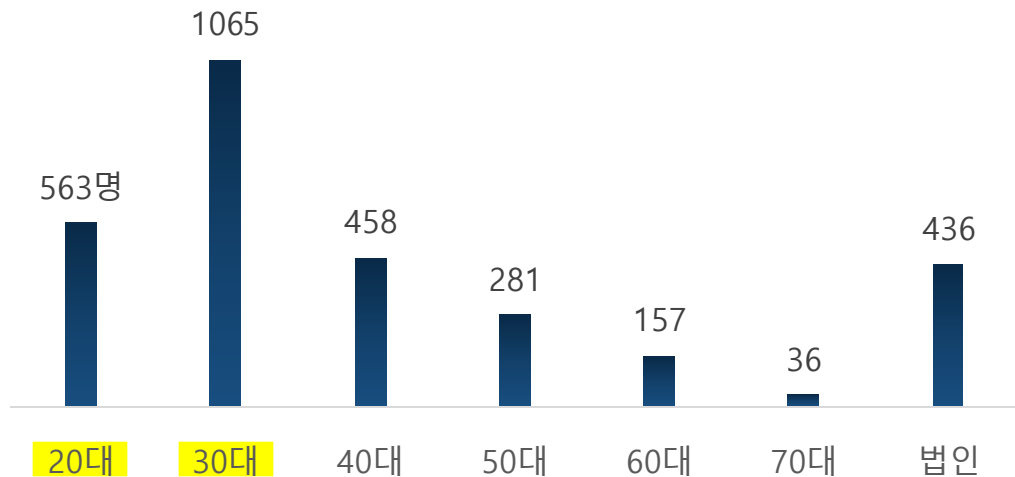
# 주제 선정 배경

## 01. 배경 소개

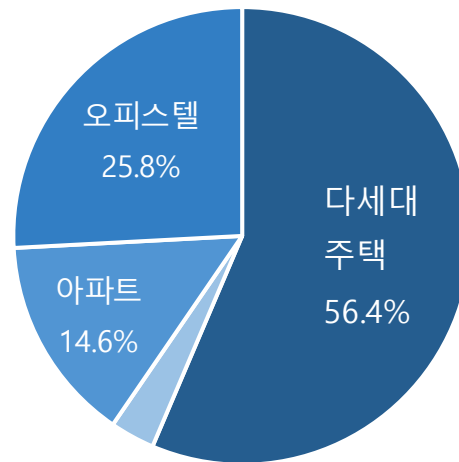
### 전세사기란?

집주인(임대인)이나 건축주, 중개인 등 전세계약 관련자들이 세입자(임차인)를 속여 전세보증금을 가로채는 행위

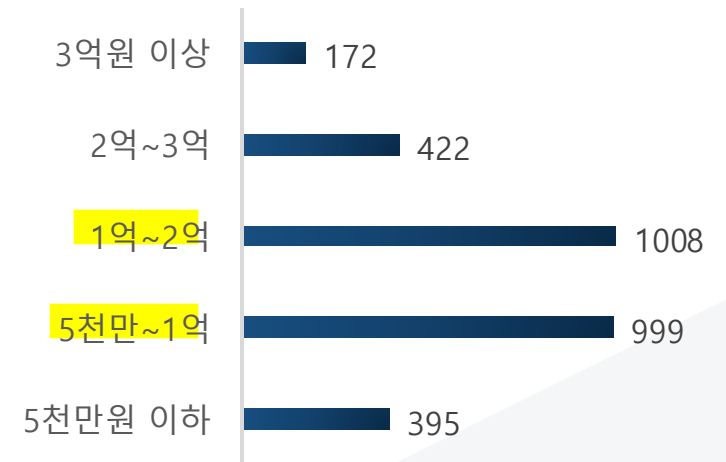
연령별



주택유형별 비중



금액별



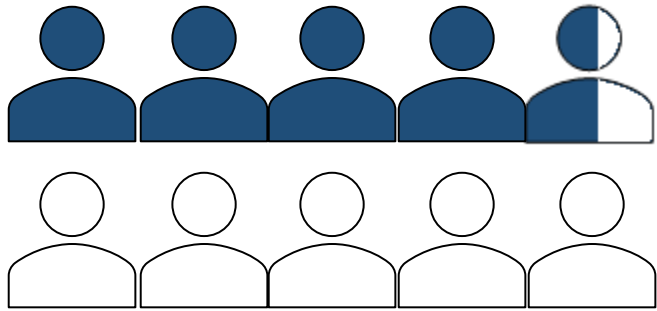
사회초년생, 서민층에 집중된 피해

# 주제 선정 배경

## 01. 배경 소개

“ 절반은 몰랐다 ”

- 빌라왕 김씨 세입자 가운데 46.5%(131명)는 김씨가 위험한 집주인인 줄 몰랐다고 답했습니다.

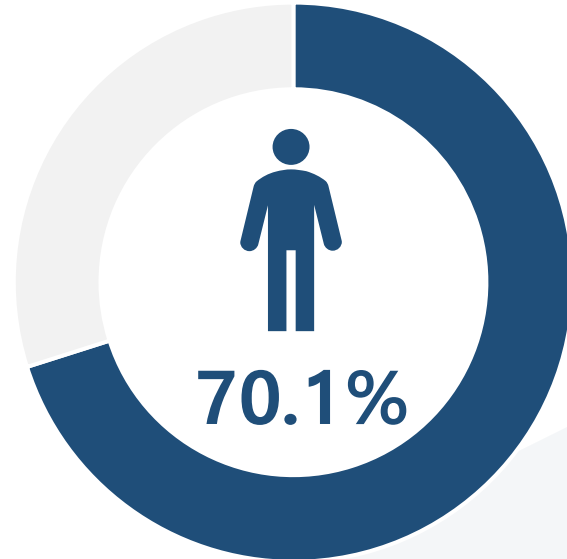


46.5%

출처 | 빌라왕 김씨 세입자 설문조사

“ 20~30대가 70% ”

- 빌라왕 사기 피해자의 70.1%(198명)가 20~30대인 것으로 나타났습니다.



출처 | MBC 기획취재팀 설문조사

✓ 전세계약 관련 **지식**이 부족할수록 전세사기 위험에 노출될 확률 증가

# 주제 선정 배경

## 02. 전세사기 대응 실태



- **전세금반환보증보험**에 가입했다라도, 보험금을 받을 수 없는 상황
  - 오히려 **보증보험을 악용**하는 전세사기 사례 증가
  - 피해자들을 보호할 수 있는 **정부 정책의 부재**
- 따라서 **전세계약 이전에 위험을 차단하는 것이 중요**

277억 가로챈 '전세 사기' 일당 구속

.. "보증보험도 무용지물"

MBC뉴스데스크 2023.03.21

"임대 계약을 한 뒤에 바뀐 집주인이 잠적하면 전세 보증 **보험에 가입을 해도 속수 무책으로 당할 수 밖에 없는 상황.**"

"오피스텔을 사고 판 집주인들간의 매매계약서나, 새 집주인의 주민번호가 있어야 보험금을 내 줄 수 있다는 겁니다."

빛에 빗 더하기...  
전세 사기 책임 피해자에 떠넘겨

KBS뉴스 2023.05.22

전세 피해 지원센터를 찾은 강 모 씨.  
신청 닷 달 만에 피해자로 인정받았습니다.  
**그러나 보증금 일부라도 돌려받기는 어렵게 됐습니다.**

**보증보험이 전세 사기에 악용**되는 것을 막지 못한 정부, 이들 조직에 쉽게 담보 대출을 내준 금융권에는  
왜 책임을 묻지 않냐는 성토도 이어집니다

임차인 올리는 '전세사기피해자법'  
사각지대

주간조선 2023.08.09

**피해자 5000명... 법 적용은 1900건 불과**

"입주 전 사기, 보증금 5억원 이상의 피해자, 밝혀진  
피해자 수가 적어 수사개시 여부가 불분명한 피해자 등 많은 피해자가 특별법 사각지대에 방치돼 있다"

"전세사기 문제는 개인이 해결할 수 없는 사회적 재난"

**정책적으로 허술한 제도상의 문제 때문에 발생**

# 주제 선정 배경

## 03. 전세사기 피해 현황

시민들의 재산과 국가 재정에 막대한 손실을 가져오는

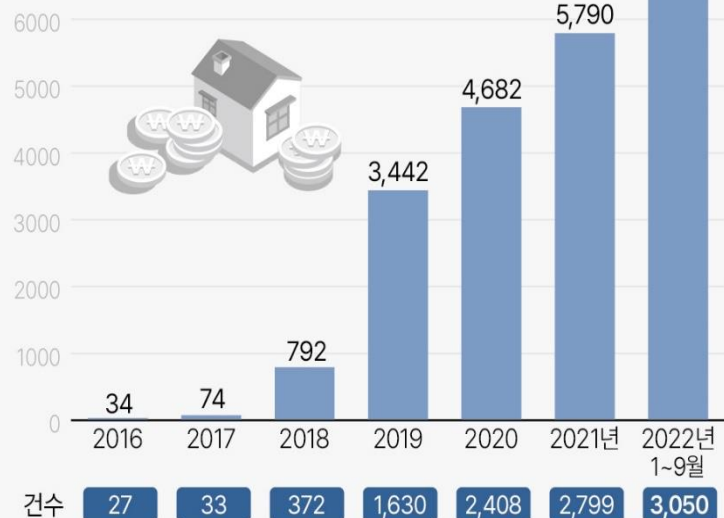
‘전세사기’ 유형 분석 및 예방 방안 제시의 필요성

### 개인 피해액 2022년까지 증가추세

#### 전세보증금 사고 금액 추이

집주인이 전세 세입자에게 보증금을 돌려주지 않은 사고 금액

단위: 억원



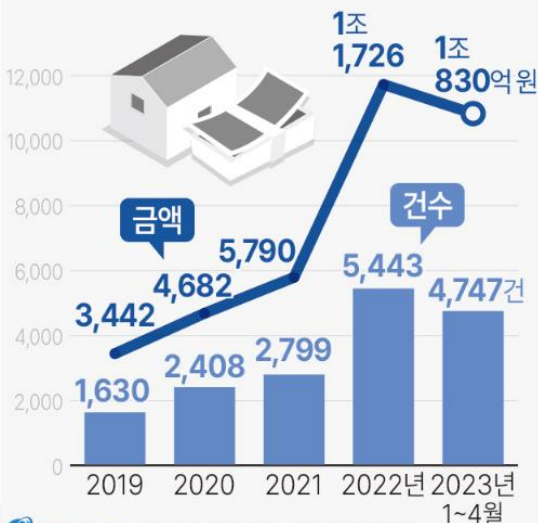
자료: 주택도시보증공사(HUG)

연합뉴스

### 2023년, 최근까지도 큰 피해

#### 전세보증사고 추이

세입자가 전세 계약해지·종료 후 1개월 내 전세보증금을 받지 못하거나 계약기간 중 경매·공매가 이뤄져 배당 후 전세보증금을 받지 못한 경우



자료: 주택도시보증공사(HUG)

### 국가 차원에서도 큰 손해

#### ■HUG 전세보증보험 추이 (단위: 억원, %)

구분	발급 규모	보증 사고	사고규모 비중	대위 변제	변제규모 비중
2015	7721	1	0.00	1	104.8
2016	5조1716	34	0.01	26	0.01
2017	9조4931	74	0.01	34	0.00
2018	19조367	792	0.04	583	0.03
2019	30조6444	3442	1.12	2836	0.92
2020	37조2595	4682	1.26	4415	1.18
2021	51조5508	5790	1.12	5040	0.97
2022	49조9280	9854	1.97	7690	1.54

# 주제 방향 및 목표

주제 설명

## 전세사기 이상탐지 모델 구축

관심 매물의 건물명 및 등기부등본 입력



건물명을 입력해주세요

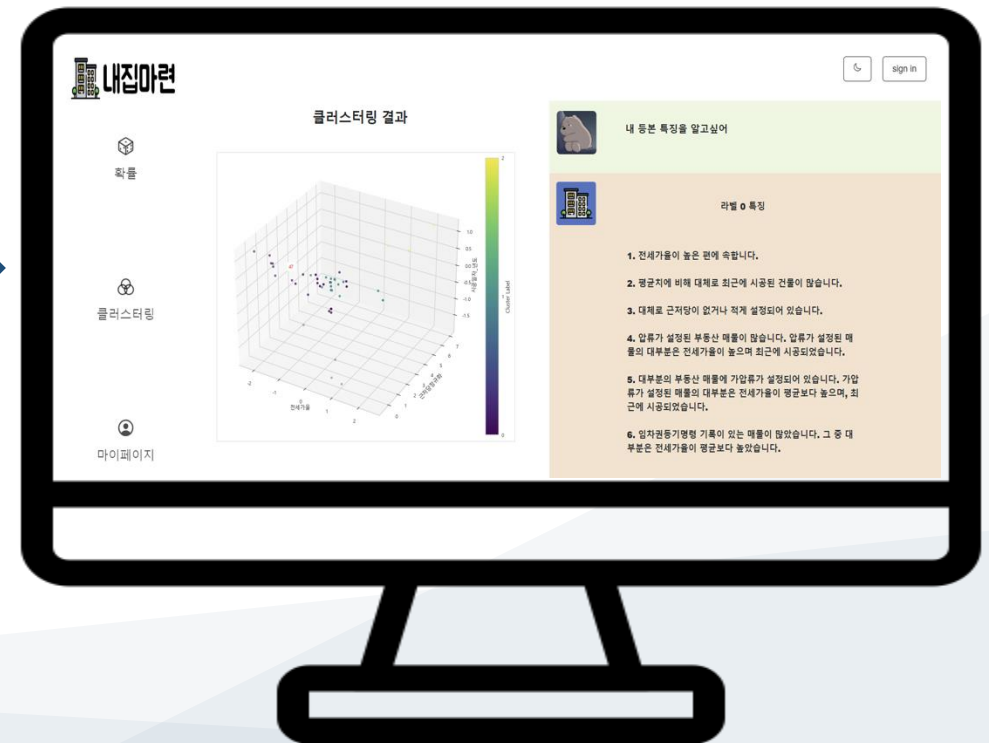
건물명

등기부등본 업로드

파일 선택 선택된 파일 없음

제출

전세사기 의심 확률 및 관련 유형&이유 제시



# 분석 방향 및 목표

선행 연구

## 선행 연구

“지역구 단위의 위험지역 선정”

“한 가지 사기 유형에 대해서만 위험도 도출”

지역 단위



강남구



서초구



양천구



## 분석 방향

“개별 부동산 매물을 대상으로 세분화”

“다양한 사기 유형에 대해 위험도 도출 가능”

개별 단위





# 분석 방향 및 목표

변수 도출 과정

## <경찰청의 전세사기 유형 분류>

- 무자본·갭투자
- 깡통전세 등 고의적 보증금 미반환
- 부동산 권리관계 허위고지
- 실소유자 행세 등 무권한 계약
- 위임범위 초과 계약
- 허위보증·보험
- 불법 중개·매개(공인중개사법위반)

## <확인할 수 있는 변수>

- 중개보조원 수, 전세가율, 시공일자
- 전세가율, 시공일자, 근저당
- 압류, 가압류 등 등기부 정보, 근저당, 중개소 정보
- 등기부의 신탁, 소유권 이전, 공동소유자 정보
- 등기부등본의 신탁, 담당 공인중개사사무소 정보
- 전세가율, 등기부 정보(전세목적, 가처분, 가압류)
- 담당 공인중개사사무소 정보(중개보조원, 공인중개사)



# 분석 방향 및 목표

활용데이터 정의 - 전세사기 피해 매물

전세사기 피해 매물 데이터 수집방법

사기데이터 라벨링 방법

- 경매 매각명세서를 통해 피해자의 전세계약 시점을 확인하였고, 첫 전세계약 시점부터 경매 개시 결정일까지의 등기부기록을 입력함

**한마음아파트 아파트**  
블로그리뷰 14

최근 극단적 선택을 한 전세사기 피해자 A씨가 살던 인천 미추홀구 송의동 353-22 한마음아파트 제7층 제703호

순위번호 등 기 목 적 접 수 등 기 원 인 권리자 및 기타사항

3 소유권이전 2017년3월8일 제79843호 2017년3월8일 신학재상의귀속 소유자 박경훈 860111-\*\*\*\*\* 인천광역시 남구 경인로 131, 2층(도파동)

2번 신학재상의귀속 신학재상의귀속

4 소유권이전 2017년3월8일 제79844호 2018년3월16일 해리 소유자 김용경 480127-\*\*\*\*\* 인천광역시 중구 반달로 13(복성동1가)

【 을 구 】 ( 소유권 이외의 권리에 관한 사항 )

순위번호 등 기 목 적 접 수 등 기 원 인 권리자 및 기타사항

1 근저당권설정 2017년3월8일 제79845호 2017년2월24일 신정제야 제권취소액 - 금141,600,000원 채무자 김용경 인천광역시 중구 반달로 13(복성동1가) 근저당권자 남원동업합동조합 211336-0000244 권리복도 남원시 중정로 100(향교동)

2 1번근저당권설정 기말소 2021년2월5일 제48134호 2021년2월5일 해리

관할등기소 인천지방법원 등기국

조사 결과 A씨는 이른바 ‘건축왕’으로 불리는 건축업자 B가 거주하고 있던 세입자로 확인됐다.

인천시에 따르면 A씨는 2018년 6월 보증금 6200만원을 냈었으나 2017년 2월 근저당이 설정된 해당 아파트는 현재

	A	C	D	E	F	G	H	I	J
1	아파트명(사기O)	전용면적 (m²)	첫 전세계약 시점	매매가_new	전세가_new	전세가율_new (=전세/매매)	임의경매 개시결정	강제경매 개시결정	시공일자
2	삼성세르빌	29.98	Sep-20	2.69	2.6	0.9665427509	2021-11-03		2020-06-26
3	혜성빌리스	48.2	Oct-20	3.3	3.3	1		2022-01-11	2008-12-19
4	딜리언하우스	43.7	Oct-20	3.55	3.52	0.9915492958	2021-11-04		2020-07-02
5	HB엘림캐슬	59.97	2021-03-30	1.83	2.48	1.355191257	2023-07-24		2019-03-11
6	신영베르디움A동 (402호)	29.9	2020-10	2.55	2.55	1		2022-12-02	2019-01-18
7	금강햇빛마을	40.8	2017-06	0.92	0.65	0.7065217391		2020-08-10	2011-09-21
8	그린아트빌	49.96	2018-12	1.07	0.85	0.7943925234	2022-07-21		2015-07-29
9	주영빌리스								
10	삼성맨션								
11	한아름 아파트								
12	청솔빌라(우이동)								
13	한집시티빌								
14	세울아파트								
15	행복타운								
16	찰북캐슬								
17	한마음 아파트								
18	새빌아파트								

K	L	M	N	O	P	Q	R	S	T	U	V
가압류	가처분	가등기	압류	소유권이전 담보가등기	소유권 이전등기	신탁	소유권 이전	임차권등기명령	공동소유자	근저당 (기간내)	근저당(정규화) new
0	0	0	0	0	0	0	0	0	0	1.95	1.691449814
0	0	0	1	0	0	0	1	1	0	0.7	1.212121212
1	0	0	1	0	0	0	0	0	0	2.25	1.625352113
0	0	0	1	0	0	0	0	1	0	2.48	2.710382514
1	0	0	1	0	1	0	0	1	0	20.4	9
1	0	0	1	0	0	0	0	0	0	0	0.7065217391
0	0	0	0	0	0	0	0	0	0	1.248	1.960747664
0	0	0	1	0	0	1	1	0	0	1.32	1.866666667
0	0	0	1	0	0	0	0	1	0	1.625	1.243315508
0	0	0	1	0	0	0	0	0	0	0.12	0.8516242318
1	0	0	1	0	1	0	0	0	0	9	5.090909091
0	0	0	0	0	0	0	0	1	0	0.6	1.382488479
0	0	0	0	0	0	0	0	0	0	1.51	1.161356629
0	0	0	0	0	1	0	0	0	0	1.34	1.523809524
1	0	0	0	0	0	0	0	0	0	1.196	1.649152542
0	0	0	0	0	0	0	0	0	0	1.416	1.404137931
0	0	0	1	0	0	0	0	0	0	1.404	1.653968254

## 활용데이터 정의 - 전세사기 아닌 매물

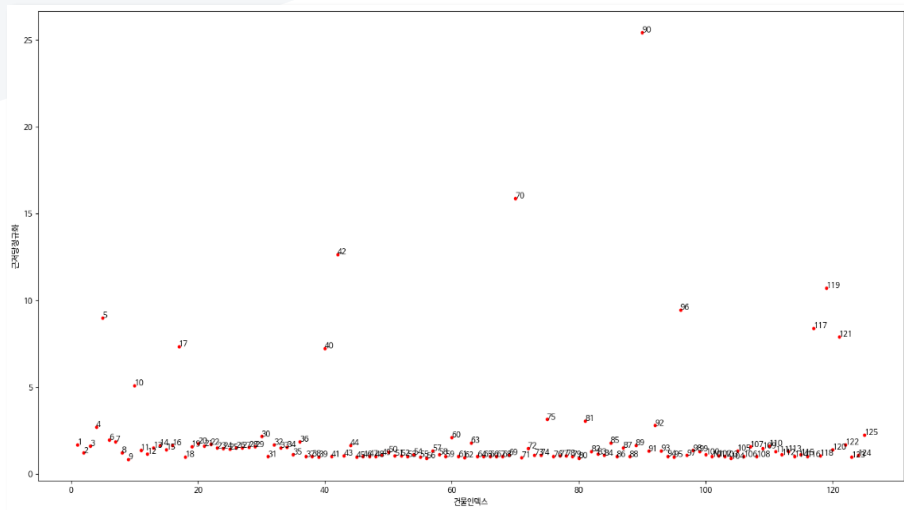
## 사기 아닌 데이터 라벨링 방법

- 첫 전세계약일부터 두 번째 전세계약일 사이의 등기부기록을 기입하였음

159	아파트명(사기X)	전용면적 (m^2)		매매가(억)	전세가(억)	전세가를	매매계약일	전세계약일	시공일자					
162	미아한일유엔아이	84			3.9	#DIV/0!		2020-02-21	2002-11-29					
163	미아한일유엔아이	84		4.13	3	0.7263922518	2018-06-05	2018-06-08	2002-11-29					
164	길음뉴타운9단지래미안	84			4.8	#DIV/0!		2016-01-25	2010-09-30					
165	길음뉴타운9단지래미안	84		5.3	5.2	0.9811320755	2015-12-03	2017-10-11	2010-09-30					
166	그린하우스	29.95		2.44	2.3	0.9426229508	2020-06-07	2020-11-14	2020-04-28					
167	그린하우스	29.95		2.53	2.2	0.8695652174	2020-06-28	2022-11-15	2020-04-28					
168	우진명가	53.98		3.35	2.95	0.8805970149	2022-01-27	2020-04-30	2019-02-08					
169	우진명가	53.98		3.35	3.05	0.9104477612	2022-01-27	2022-03-31	2019-02-08					
170	궁궐별장빌		가압류	가처분	가등기	압류	소유권이전 담보가등기	소유권이전 이전등기	신탁	소유권이전	임차권등기명령	공동소유자	근저당	근저당(정규화) new
171	궁궐별장빌													
172	정원부르네													
173	나산아파트	0	0	0	0	0	0	0	0	0	0	0	0.7263922518	
174	나산아파트													
175	대동아파트	0	0	0	0	0	0	0	0	0	0	0.6	1.094339623	
176	대동아파트	0	0	0	0	0	0	0	0	0	0	1.62	1.606557377	
177	남부아주													
178	남부아주	0	0	0	0	0	1	0	0	0	0	0	0.8805970149	
179	보리미즌빌													
180	보리미즌빌	0	0	0	0	0	0	0	0	0	0	0	1.05	
181	아주													
		0	0	0	0	0	0	0	0	0	0	0	0.7619047619	
		0	0	0	0	0	0	0	0	0	0	0	0.4583333333	
		0	0	0	0	0	0	0	0	0	0	0	0.8620689655	
		0	0	0	0	0	0	0	0	0	0	0.36	0.9830508475	

# 전처리

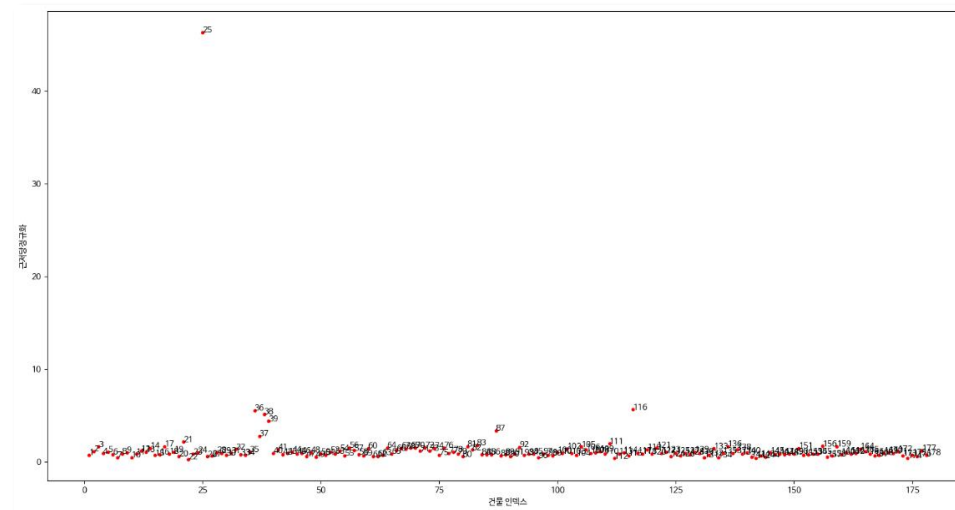
## 전세사기 데이터



index	아파트명(사기 O)	매매가_new	전세가_new	전세가율	시공일자	가압류	압류	신탁	소유권이전	임차권등기명령	근저당	근저당정규화	시공일자_datetime	시공일자_년도
20	20 토브하우스 (304호)	1.13	1.99	1.761062	2018-09-05	0	0	0	1	1	0.0000	1.761062	2018-09-05	2018
30	30 토브하우스 (502호)	1.18	2.55	2.161017	2018-09-23	0	1	1	1	1	0.0000	2.161017	2018-09-23	2018
87	87 동방빌라	0.55	0.25	0.454545	1991-05-20	1	1	0	1	0	0.5775	1.504545	1991-05-20	1991
89	89 신의웨스트빌	1.40	2.30	1.642857	2016-09-21	1	1	0	0	1	0.0000	1.642857	2016-09-21	2016

✓ Data 양 부족으로 의미있는 데이터로 남겨두기로 결정

## 전세사기 아닌 데이터



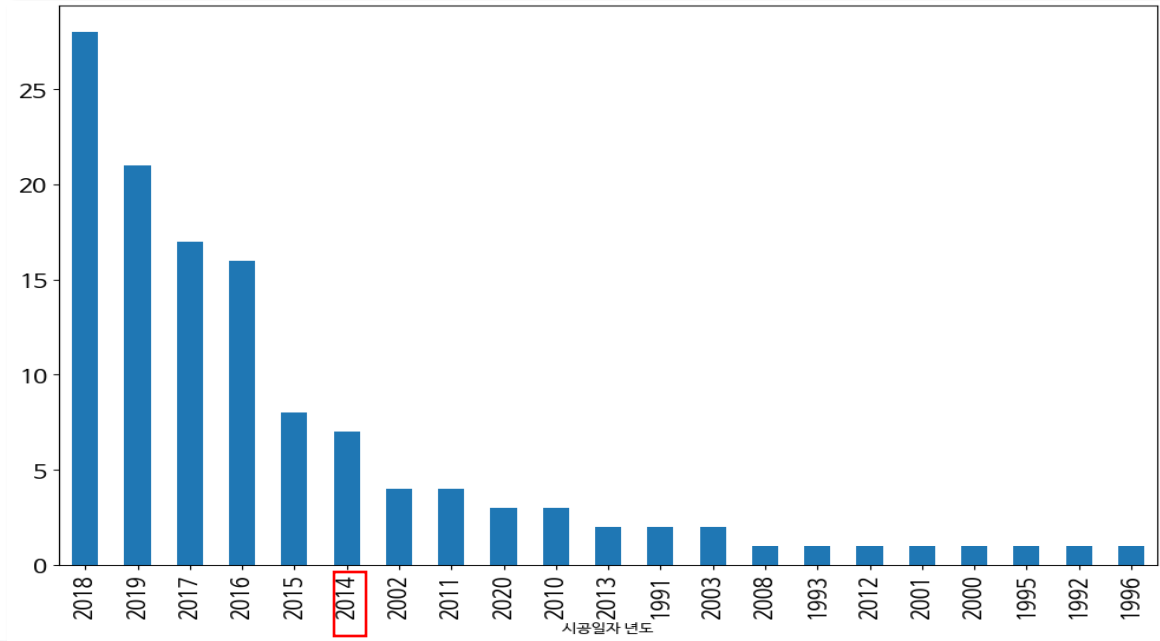
index	아파트명(사기 X)	매매가 (억)	전세가 (억)	전세가율	시공일자	가압류	압류	신탁	소유권이전	임차권등기명령	근저당	근저당정규화	시공일자_datetime	시공일자_년도
25	25 문현프라다	0.9000	0.9	1.000000	2020-10-20	0.0	0.0	0.0	0.0	0.0	40.800	46.333333	2020-10-20	2020
36	36 골든패밀리	0.4915	0.4	0.813835	2012-09-05	0.0	0.0	0.0	0.0	0.0	2.316	5.525941	2012-09-05	2012
38	38 동현빌	1.2700	1.3	1.023622	2003-06-04	0.0	0.0	0.0	0.0	0.0	5.200	5.118110	2003-06-04	2003
39	39 다모아하우스	0.9700	0.8	0.824742	2012-01-13	0.0	0.0	0.0	0.0	0.0	3.432	4.362887	2012-01-13	2012
116	116 레오파레스	1.9200	1.8	0.937500	2014-12-09	0.0	0.0	0.0	0.0	0.0	9.000	5.625000	2014-12-09	2014

✓ 등기부등본에 공동담보목록이 없는 상태  
=> 개별 근저당을 구할 수 없는 데이터

'잘못된 데이터이기에 제거'

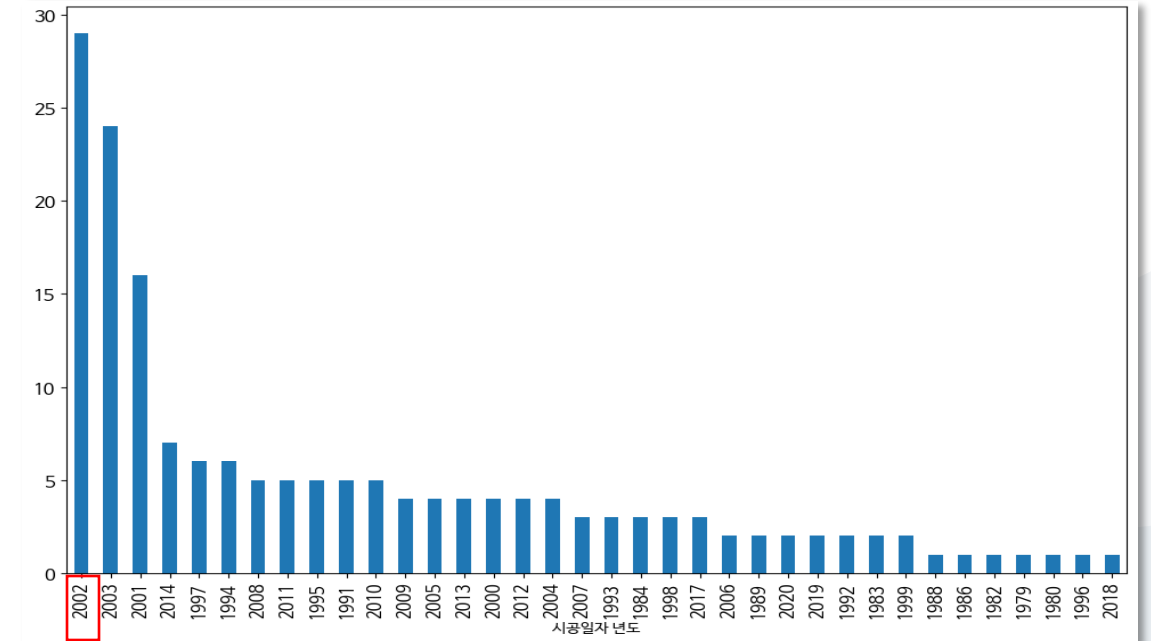
## 전세사기 데이터

	매매가(억)	전세가(억)	전세가율	가압류	압류	신탁	소유권이전	임차권등기명령	근저당	근저당정규화	시공일자_년도	label
count	125.000000	125.000000	125.000000	125.000000	125.000000	125.000000	125.000000	125.000000	125.000000	125.000000	125.000000	125.0
mean	1.916584	1.858960	0.960708	0.520000	0.656000	0.024000	0.112000	0.624000	2.391166	2.165323	2014.552000	1.0
std	0.588376	0.737486	0.269139	0.501610	0.476953	0.153665	0.316636	0.486329	7.025517	3.171365	6.503666	0.0
min	0.550000	0.250000	0.374332	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.851624	1991.000000	1.0
25%	1.500000	1.420000	0.847518	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	2014.000000	1.0
50%	1.900000	1.990000	1.000000	1.000000	1.000000	0.000000	0.000000	1.000000	0.120000	1.279762	2017.000000	1.0
75%	2.240000	2.300000	1.066667	1.000000	1.000000	0.000000	0.000000	1.000000	1.340000	1.642857	2018.000000	1.0
max	3.600000	3.600000	2.161017	1.000000	1.000000	1.000000	1.000000	1.000000	60.000000	25.448980	2020.000000	1.0



## 전세사기 아닌 데이터

	매매가(억)	전세가(억)	전세가율	가압류	압류	신탁	소유권이전	임차권등기명령	근저당	근저당정규화	시공일자_년도	label
count	173.000000	173.000000	173.000000	173.000000	173.000000	173.000000	173.000000	173.000000	173.000000	173.000000	173.000000	173.0
mean	1.858584	1.425838	0.772267	0.040462	0.046243	0.017341	0.028902	0.028902	0.345110	0.968951	2002.190751	0.0
std	0.846873	0.735510	0.202100	0.197613	0.210620	0.130918	0.168017	0.168017	0.647218	0.408503	8.063954	0.0
min	0.285000	0.190000	0.171429	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.266667	1979.000000	0.0
25%	1.330000	1.000000	0.668693	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.726392	1999.000000	0.0
50%	1.750000	1.300000	0.809524	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.872727	2002.000000	0.0
75%	2.220000	1.800000	0.894737	0.000000	0.000000	0.000000	0.000000	0.000000	0.600000	1.111111	2007.000000	0.0
max	6.500000	5.200000	1.428571	1.000000	1.000000	1.000000	1.000000	1.000000	3.960000	3.357143	2020.000000	0.0



✓ 전세사기 : 전세가율 평균 ↑, 가압류 키워드 확률 ↑, 신탁 ↑, 근저당 ↑

# 모델 설명

## 데이터 가공

	A	B	C	D	E	G	H	I	J	K	L	M	N	P	Q
1		아파트명	매매가(억)	전세가(억)	전세가율	가압류	압류	신탁	소유권이전	임차권등기	근저당	근저당정기	지역	시공일자	label
2	0	삼성쉐르빌	2.69	2.6	0.966543	0	0	0	0	0	1.95	1.69145	서울 강서	2020	1
3	1	혜성팰리스	3.3	3.3	1	0	1	0	1	1	0.7	1.212121	서울 강서	2008	1
4	2	밀리언하우스	3.55	3.52	0.991549	1	1	0	0	0	2.25	1.625352	서울 중랑	2020	1
5	3	HB엘림캐슬	1.83	2.48	1.355191	0	1	0	0	1	2.48	2.710383	인천 부평	2019	1
6	4	신영베르디	2.55	2.55	1	1	1	0	0	1	20.4	9	서울 강서	2019	1
7	5	그린아트빌	1.07	0.85	0.794393	0	0	0	0	0	1.248	1.960748	인천 미추	2015	1
8	6	주영팰리스	1.35	1.2	0.888889	0	1	1	1	0	1.32	1.866667	인천 미추	2014	1
9	7	삼성맨션	1.87	0.7	0.374332	0	1	0	0	1	1.625	1.243316	인천 미추	2010	1
10	8	한아름 아파트	1.139	0.85	0.746269	0	1	0	0	0	0.12	0.851624	인천 미추	2015	1
11	9	청솔빌라(아파트)	2.2	2.2	1	1	1	0	0	0	9				
12	10	한집시티빌	0.868	0.6	0.691244	0	0	0	0	1	0.6				
13	11	세움아파트	1.946	0.75	0.385406	0	0	0	0	0	1.51				
14	12	행복타운	1.26	0.58	0.460317	0	0	0	0	0	1.34				
15	13	행복캐슬	1.18	0.75	0.635593	1	0	0	0	0	1.196				
16	14	한마음 아파트	1.45	0.62	0.427586	0	0	0	0	0	1.416				
17	15	새힘아파트	1.26	0.68	0.539683	0	1	0	0	0	1.404	1.653968	인천광역시	2016	1

0      173  
1      125  
Name: label, dtype: int64

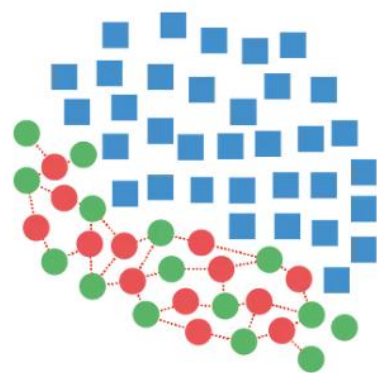
✓ imbalanced data & Insufficient data

## How to solve?

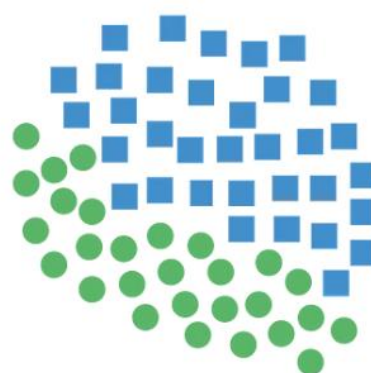
### Synthetic Minority Oversampling Technique



Original Dataset

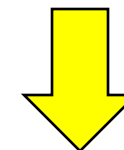


Generating Samples



Resampled Dataset

```
0    173  
1    125  
Name: label, dtype: int64
```



```
0    180  
1    150  
Name: label, dtype: int64
```



## GridsearchCV

- ✓ 0. 모델 성능 평가 지표, 교차 검증
- ✓ 1. Logistic Regression
- ✓ 2. **K-Nearest** Neighbor (**KNN**)
- ✓ 3. Support Vector Classifier (SVC)
- ✓ 4. Decision Tree Classifier



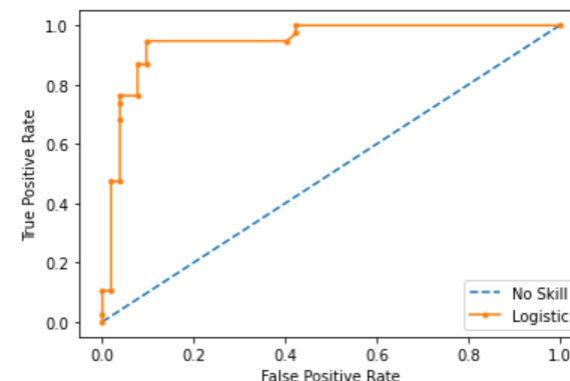
### WHAT TO SEE?

실제 0 (음성 클래스)	<b>TN (True Negative)</b> 0을 0으로 잘 예측한 경우	<b>FP (False Positive)</b> 0을 1로 잘못 예측한 경우
실제 1 (양성 클래스)	<b>FN (False Negative)</b> 1을 0으로 잘못 예측한 경우	<b>TP (True Positive)</b> 1을 1로 잘 예측한 경우
코딩재개발	예측 0 (음성 예측)	예측 1 (양성 예측)

### ➤ Recall & AUC score

$$Recall = \frac{TP}{TP + FN}$$

No Skill: ROC AUC=0.500  
Logistic Regression: ROC AUC=0.944

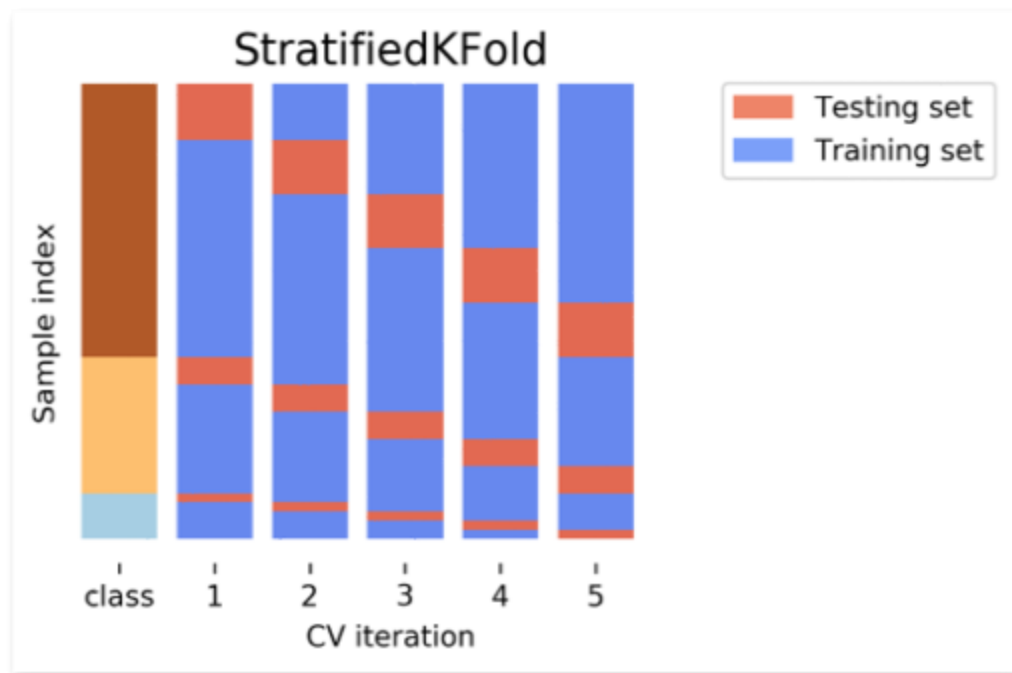


✓ 실제 전세 사기인 것을 전세 사기라고 판단하는 능력이 모델의 핵심!!!

# GridsearchCV

## 0. 교차 검증

### Mission : 공정한 상황에서 검증



교차검증 : 1 번째  
학습 데이터의 레이블 분포 :  
0 144  
1 120  
Name: label, dtype: int64

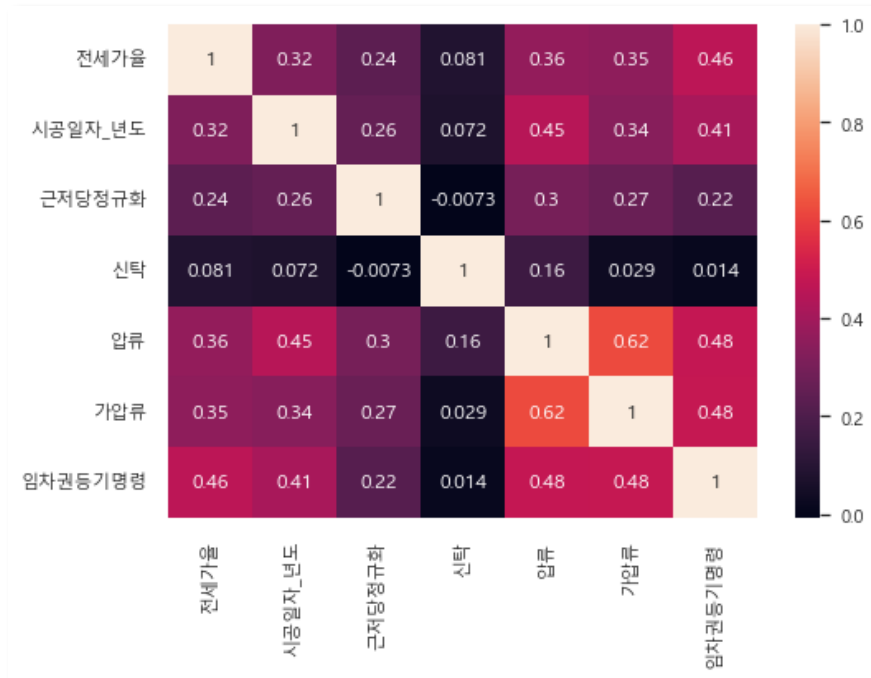
검증 데이터의 레이블 분포 :  
0 36  
1 30  
Name: label, dtype: int64

교차검증 : 2 번째  
학습 데이터의 레이블 분포 :  
0 144  
1 120  
Name: label, dtype: int64

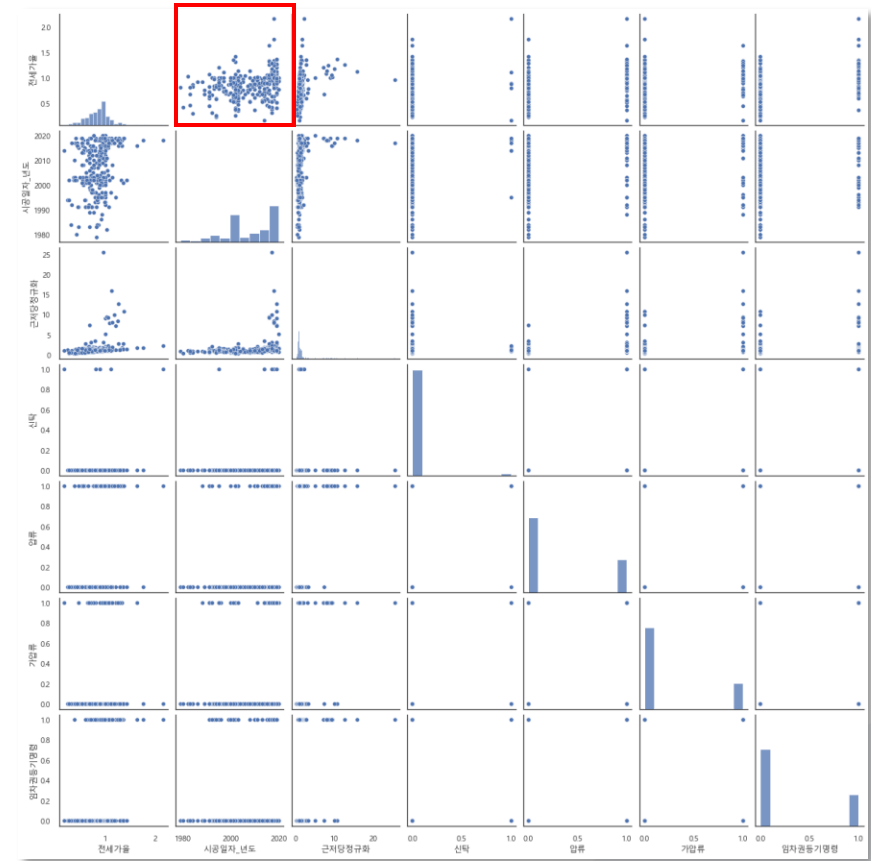
검증 데이터의 레이블 분포 :  
0 36  
1 30  
Name: label, dtype: int64

# GridsearchCV

## 1. Logistic Regression



상관계수 - 시각화



산점도를 통해 선형 관계 확인

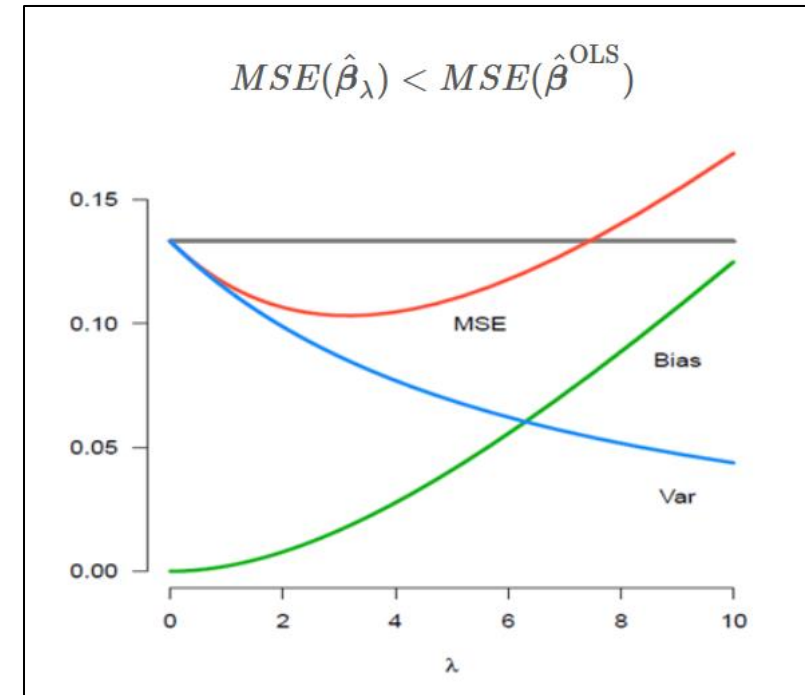
# GridsearchCV

## 1. Logistic Regression

	VIF Factor	features
0	17.875386	전세가율
1	14.692437	시공일자_년도
2	1.675790	근저당정규화
3	1.061241	신탁
4	2.689489	압류
5	2.367197	가압류
6	2.192125	임차권등기명령

다중공선성 : 회귀 분석의 기본 가정  
위배

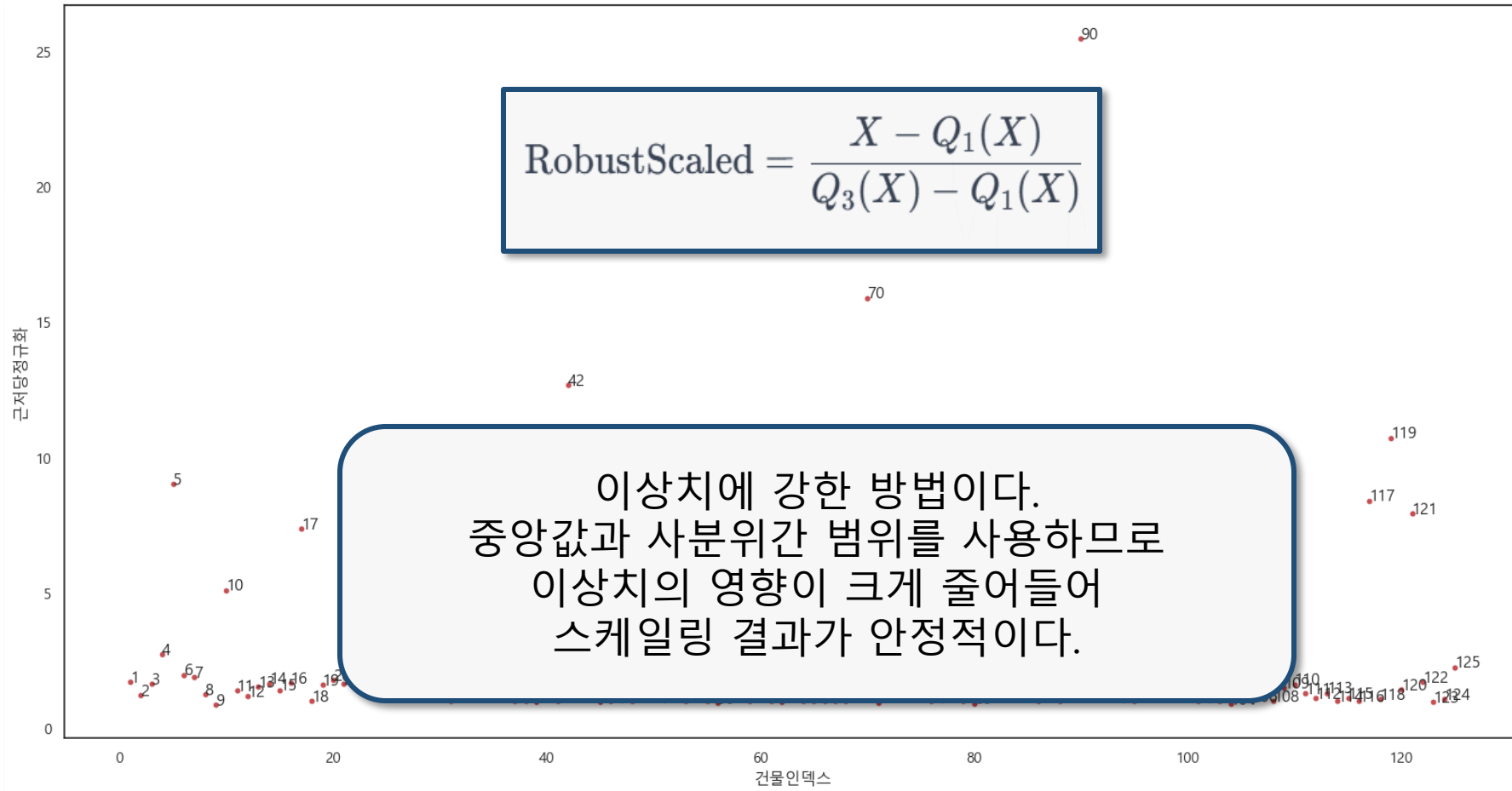
## 'l2' regularization (Ridge)



⇒ GridsearchCV 에서 파라미터로 'L1', 'L2' Regularization 을 주어 이를 해결하고자 함

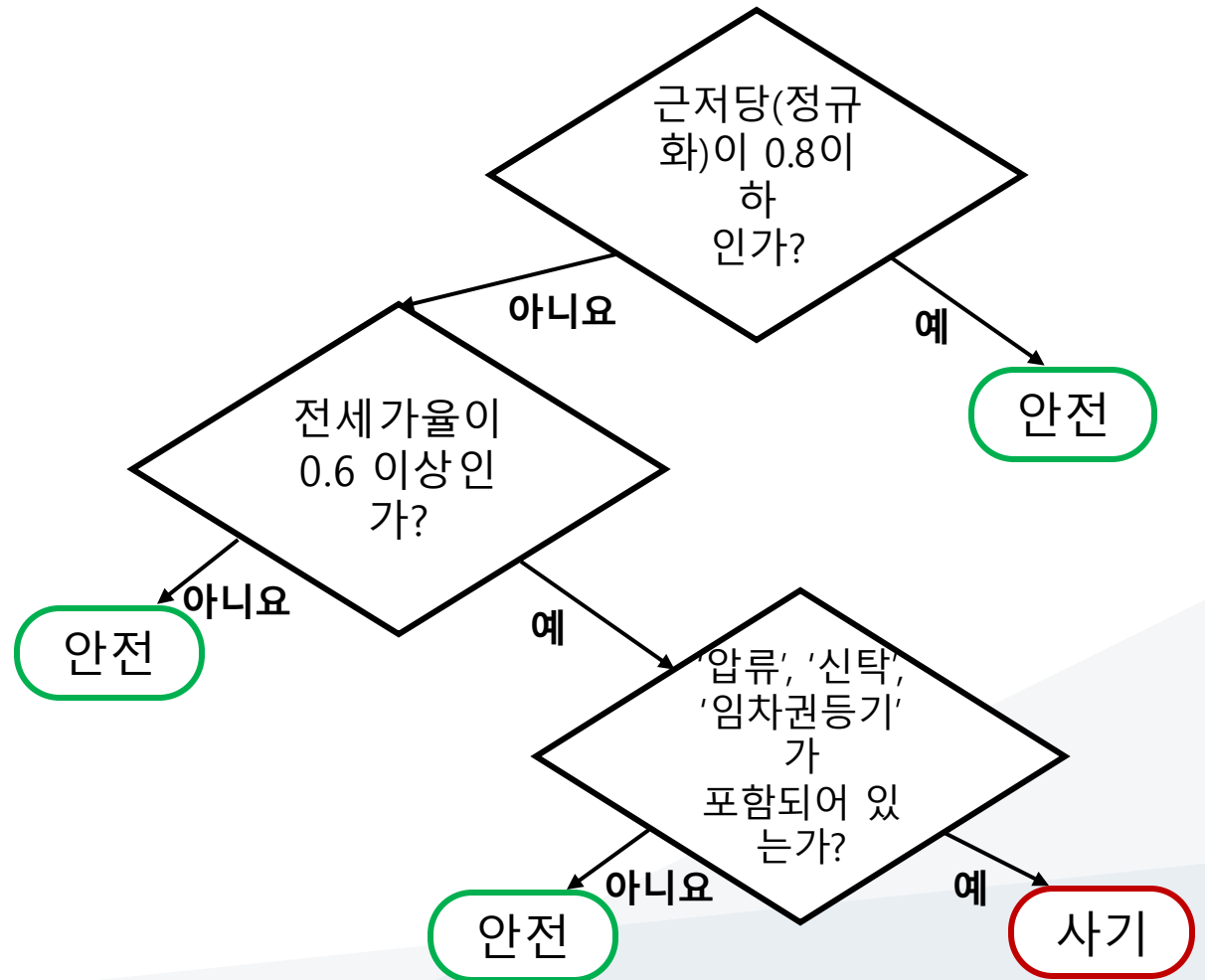
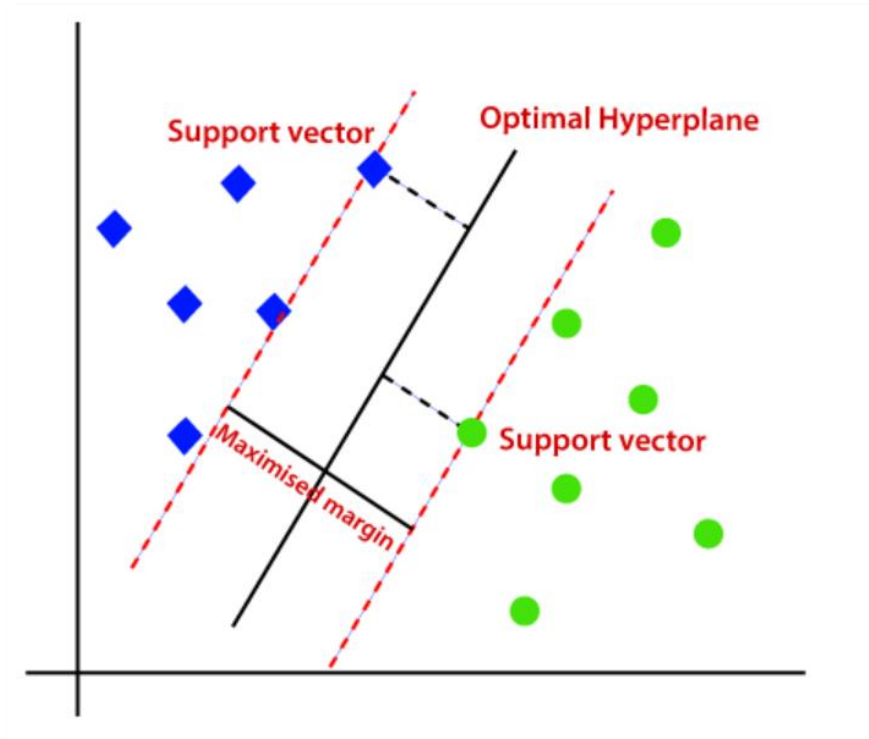
# GridsearchCV

## 2. K-Nearest Neighbor (KNN)

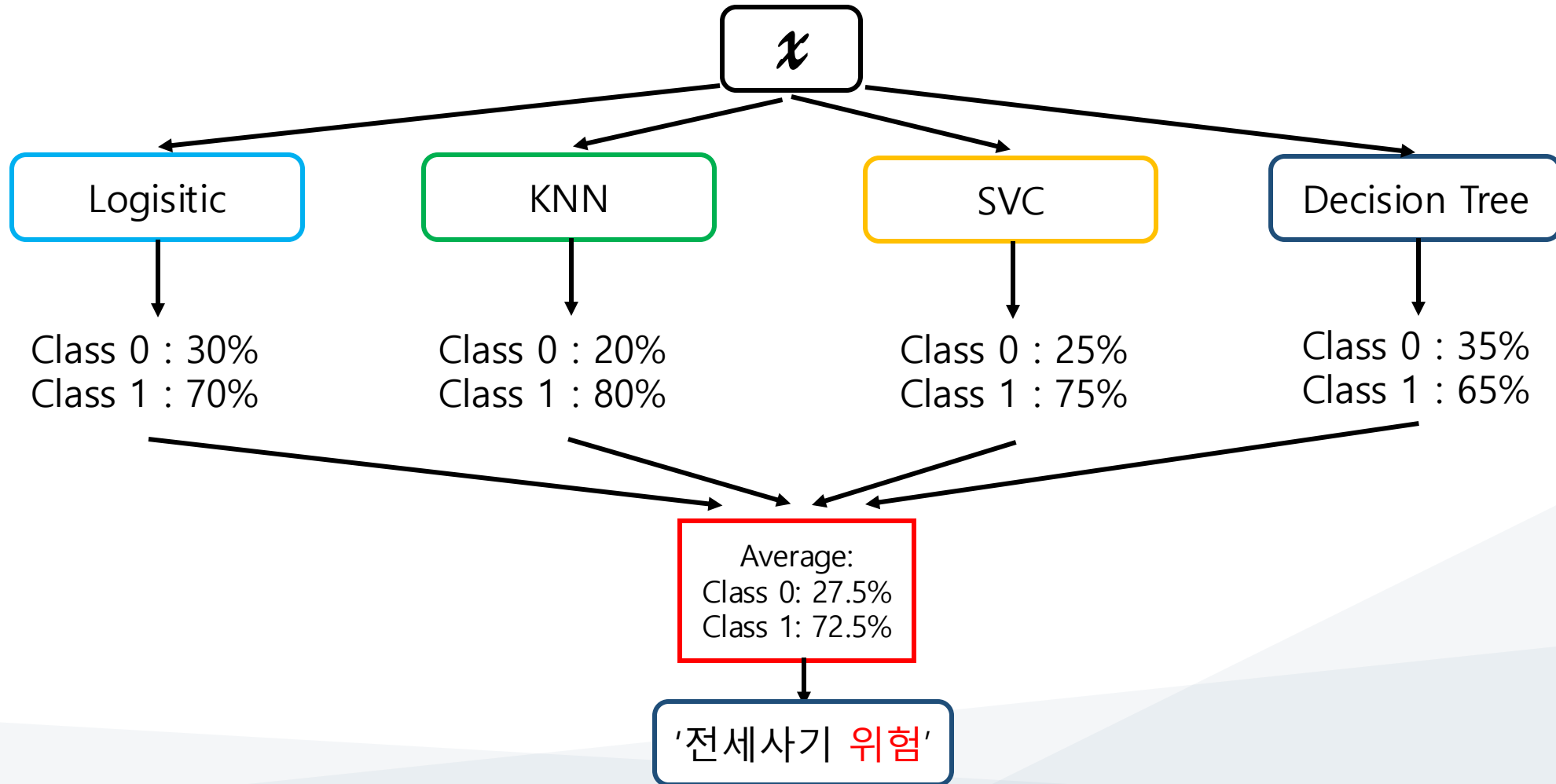


# GridsearchCV

## 3. Support Vector Classifier (SVC) & 4. Decision Tree Classifier



### Soft Voting



# GridsearchCV

## Soft Voting

## Soft Voting

Logistic

Lreg 파라미터: {'C': 1, 'penalty': 'l2'}  
Lreg 최고점수: 0.8566

실제값과 예측값 정확도 auc: 0.8814814814815  
실제값과 예측값 정확도 recall: 0.8  
실제값과 예측값 정확도 f1: 0.8674698795180723

KNN

KNN 파라미터: {'algorithm': 'auto', 'metric': 'manhattan', 'n\_ne  
KNN 예측 정확도: 0.9287  
KNN 파라미터: {'algorithm': 'auto', 'metric': 'manhattan', 'n\_ne  
KNN 예측 정확도: 0.9300

실제값과 예측값 정확도 auc score: 0.874074074074074  
실제값과 예측값 정확도 recall: 0.822222222222222  
실제값과 예측값 정확도 f1: 0.8604651162790697

SVC

SVC 파라미터: {'C': 10, 'break\_ties': False, 'gamma': 'auto'  
SVC 최고점수: 0.8816

실제값과 예측값 정확도 auc score: 0.8962962962962964  
실제값과 예측값 정확도 recall: 0.8666666666666667  
실제값과 예측값 정확도 f1: 0.8863636363636364

Decision Tree

decisiontree 파라미터: {'ccp\_alpha': 0, 'criterion': 'gini', 'max\_feature  
decisiontree 예측 정확도: 0.8970  
decisiontree 파라미터: {'ccp\_alpha': 0, 'criterion': 'entropy', 'max\_feat  
decisiontree 예측 정확도: 0.9045

실제값과 예측값 정확도 auc score: 0.9111111111111111  
실제값과 예측값 정확도 recall: 0.9333333333333333  
실제값과 예측값 정확도 f1: 0.9032258064516129

# 여러 성능지표 확인

```
pred_svot = svot_clf.predict(X_test)
print(f"실제값과 예측값 정확도 auc score: {roc_auc_score(y_test, pred_svot)}")
print(f"실제값과 예측값 정확도 recall: {recall_score(y_test, pred_svot)}")
print(f"실제값과 예측값 정확도 f1: {f1_score(y_test, pred_svot)}")
```

실제값과 예측값 정확도 auc score: 0.9111111111111111  
실제값과 예측값 정확도 recall: 0.9333333333333333  
실제값과 예측값 정확도 f1: 0.9032258064516129



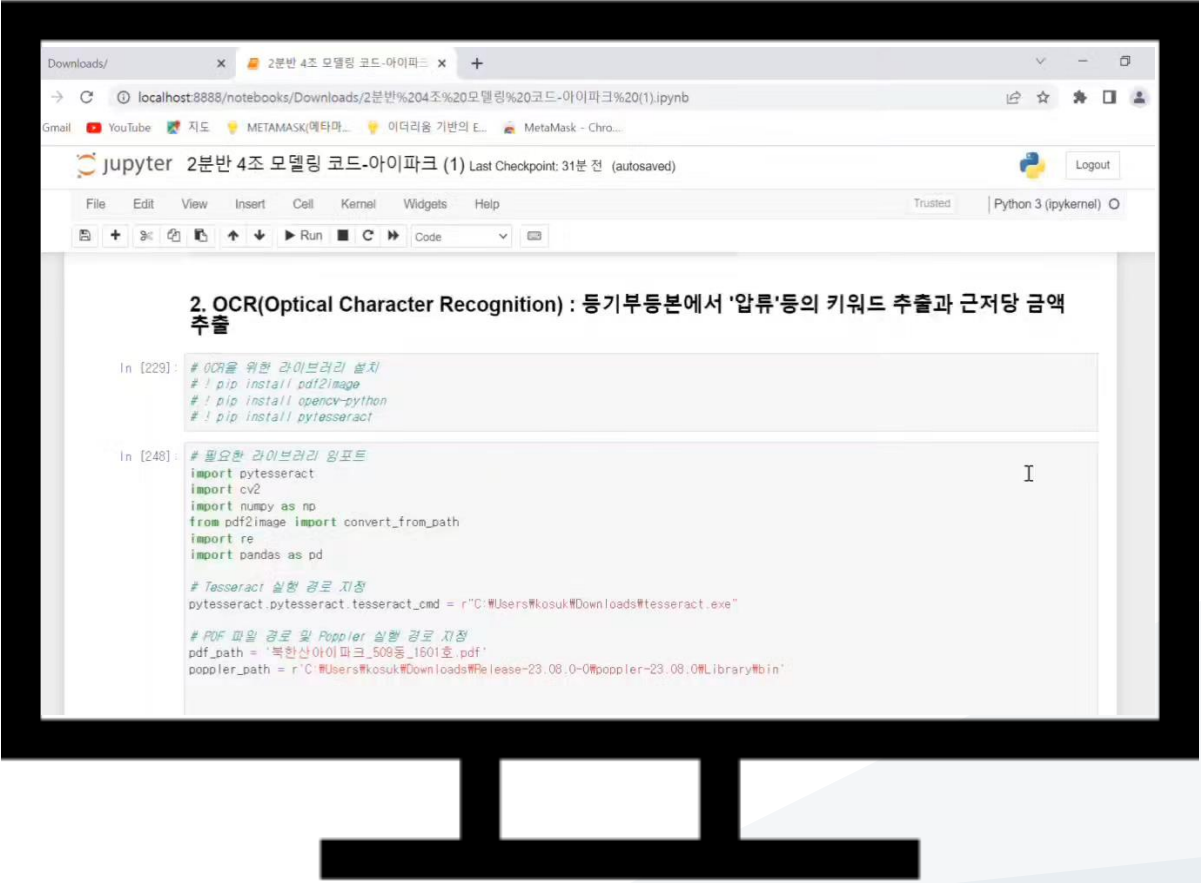
# 새로운 매물 찾기

## 크롤링



아파트명	매매가(억)	전세가(억)	시공일자
0 북한산아이파크	11억4,000	5억5,000	2004.07.30

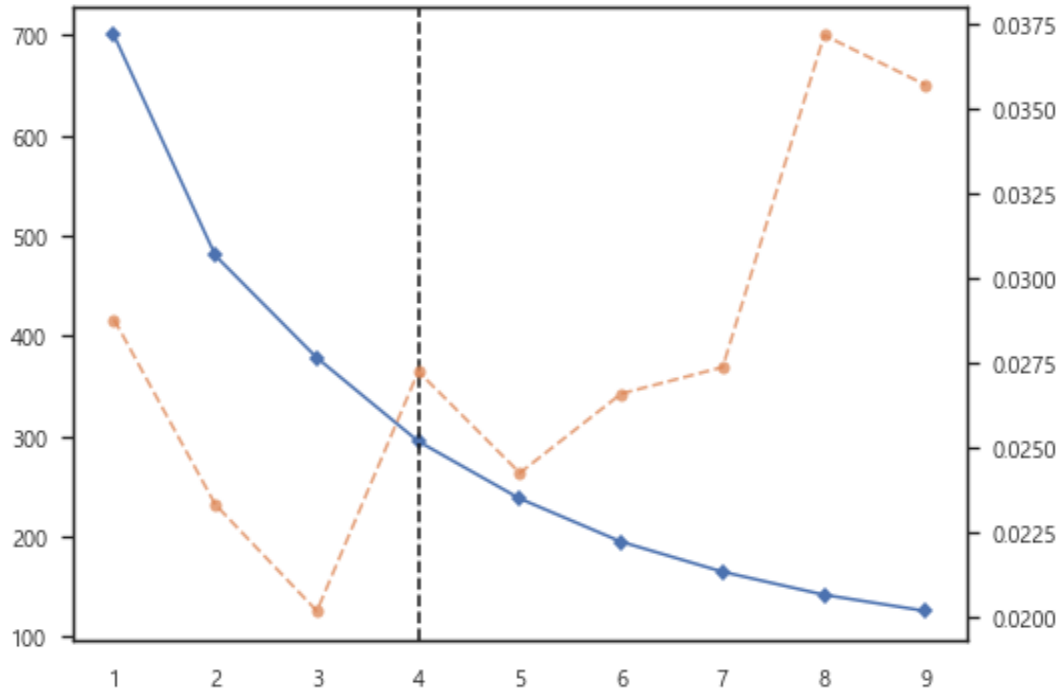
## OCR



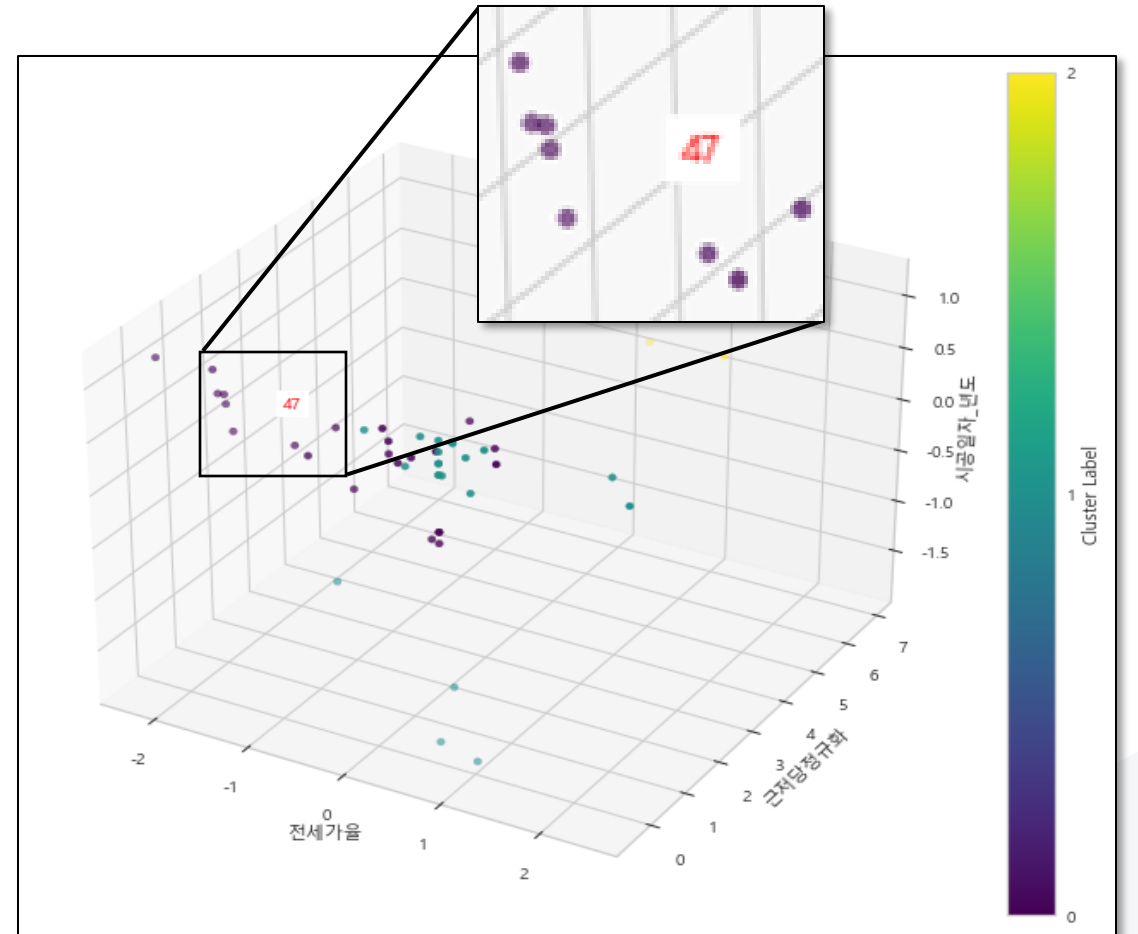
전세가율	시공일자_년도	근저당정규화	신탁	압류	가압류	임차권등기명령
0 0.482456	2004	0.607895	0	0	0	0

# 클러스터링

## Elbow Method&visualization



$k = 3 ? k = 4 ?$



전세가을	시공일자_년도	근저당정규화	신탁	압류	가압류	임차권등기명령	label
-0.750335	1.130718	-0.426701	-0.100504	-0.639101	-0.51558	-0.608164	0.0

## 결과 도출

주소: 서울특별시 도봉구 창동 825 북한산아이파크 제509동

사기일 확률은 : 10.30% 입니다

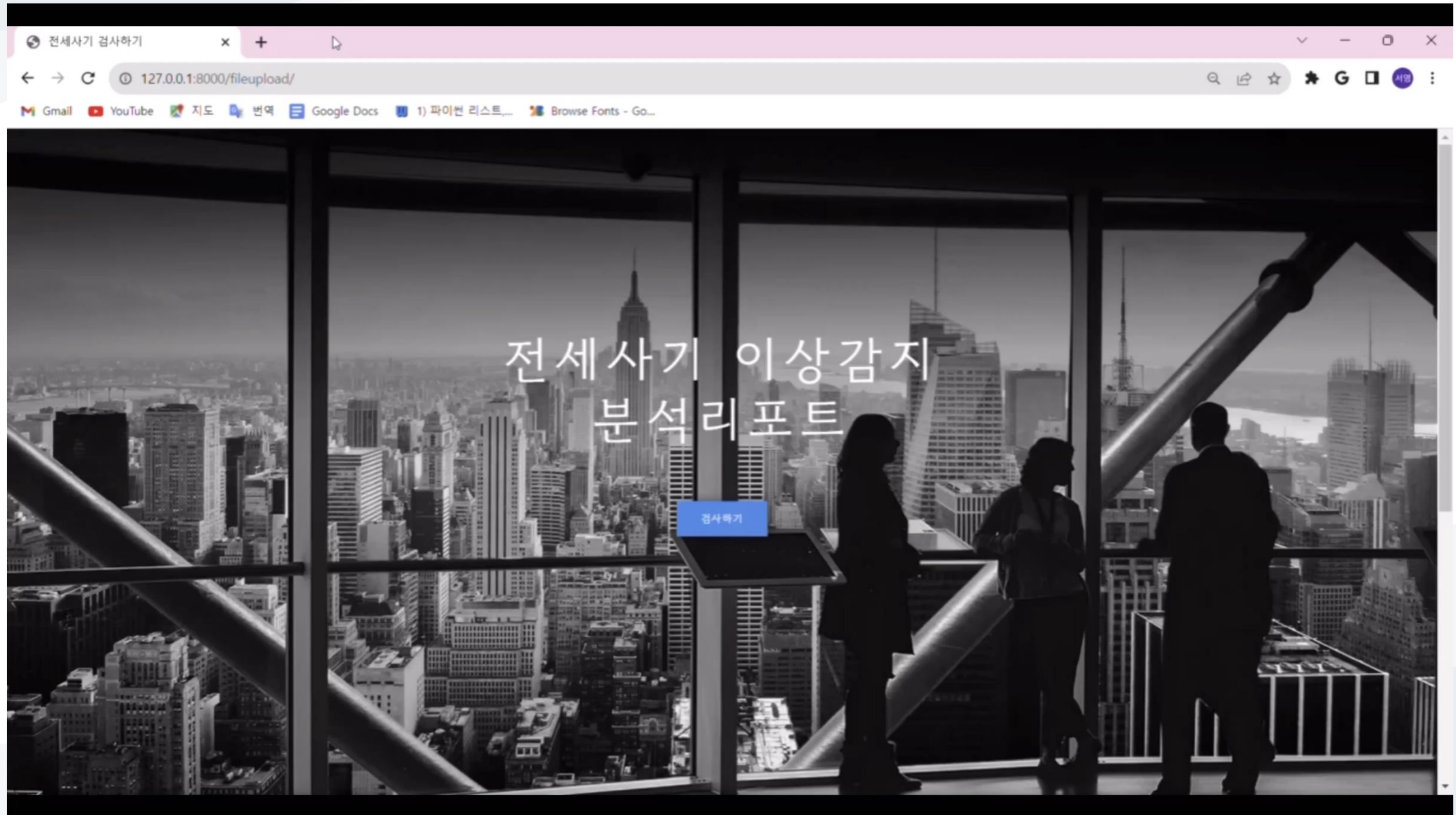
전세가율	시공일자_년도	근저당정규화	신탁	압류	가압류	임차권등기명령
0 0.495146	2004	0.633981	0	0	0	0

주소: 경상남도 김해시 삼계동 1545 삼계서희스타힐스 제108동

사기일 확률은 : 35.53% 입니다

전세가율	시공일자_년도	근저당정규화	신탁	압류	가압류	임차권등기명령
0 0.704	2019	0.704	0	0	0	0

# 웹 시연 영상



# 클러스터링 결론 도출



이 집 전세계약하기에 문제가 있을까?



## (의심되는 전세사기 유형)

### 1. 무자본 갭투자 사기

#### 정의 :

신축빌라 등 건축주, 공인중개사·분양대행사 등 브로커, 매수인 등이 공모, 미분양 빌라 등을 무자본 매입 후 임대차계약으로 보증금을 편취하는 사기 유형

#### 선정 이유:

- **최근 시공된 건물** : 신축빌라는 정확한 시세 산정이 어렵기 때문에, 전세금이 공시가격의 150%에 달해도 보증보험이 가입된다는 점을 악용하여 세입자를 안심시킨 후 매매가보다 높은 전세금을 받고 잠적하는 경우가 있습니다.

- **높은 전세가율** : 해당 매물이 경매로 넘어갈 경우, 유찰 1회당 평균적으로 경매가의 10%씩 줄어들기 때문에, 우선순위가 있더라도 낙찰된 금액이 낮기 때문에 전세금을 온전히 받을 수 없기 때문입니다.

- **근저당이 없거나 적게 설정되어 있어도 위험** : 분석 결과 근저당이 설정되어 있지 않더라도 피해사례가 많았습니다.



이 집 전세계약하기에 문제가 있을까?



## (예방 방안)

- **임대인 채무기록 확인** : 임대인의 보유 건물 수와 대출 내역을 확인하여 위험한 집주인 여부를 판단하세요

- **미분양여부 및 거래회전율 확인** : 시공일자 최근일 경우 미분양된 호수가 많다면 전세보증보험 가입이 되더라도, 보증금을 돌려 받지 못할 가능성이 있기 때문에, 해당 건물의 부동산 거래회전율을 파악하여 전세거래가 활성화되고 있는지 확인이 필요합니다.

- **전세금반환보증보험 가입** : 등기부기록에 이상이 없는 매물에 주로 발생하는 사기 유형이므로 전세금반환보증에 가입하여 예상치 못한 위험에 대비하는 것을 추천드립니다.

# 결론

## 서비스 활용방안 & 기대효과

### 서비스 활용방안

전세계약을 새로 하려는  
사람들에게 서비스를  
무료로 제공하여  
전세사기 피해 예방에 기여

### 개인적 측면

- **개인의 실수**를 줄여줄 수 있음 (등기부를 잘 못 봐서 생기는 문제 예방 가능)
- 전세사기와 관련된 전반적인 지식이 아닌, 사용자가 전세 계약을 하는 데 있어 **관련 있는 등기부정보만 추출**하여 제공하기 때문에 더욱 예방에 효과적
- 현재 등기부등본 내역 뿐만 아니라 말소 내역까지 꼼꼼히 살펴서 어떤 위험상황들이 예상되는지 제공하기 때문에 더욱 유용함

### 기대효과

### 공익적 측면

- 전세사기 예방을 통해 부동산 시장의 신뢰도 및 사회적 신뢰 향상에 기여
- 시민들의 부동산 거래에 대한 안정감 증진

# 감사합니다



2023.08.31