

# 나홀로 소송을 준비하는 일반인을 위한 법률 정보 검색 시스템 구축

팀 빈센조 | 지승빈 박병현 이찬규 임채균 지유하

---

# **CONTENTS**

---

01) 문제 정의 및 연구 배경

02) 연구 목표

03) 연구 과정

04) 연구 결과

05) 토의, 한계 및 미래연구

06) 역할 분담

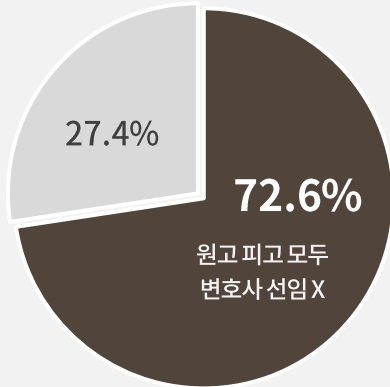
07) References



# 1. 문제 정의 및 연구 배경

## 1) 문제 정의

### “늘어나는 나 홀로 소송” <매일경제>

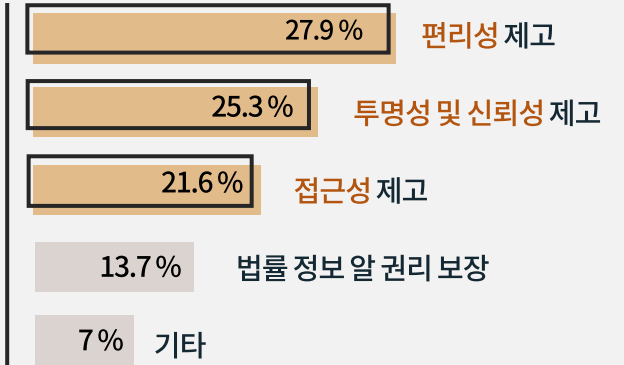
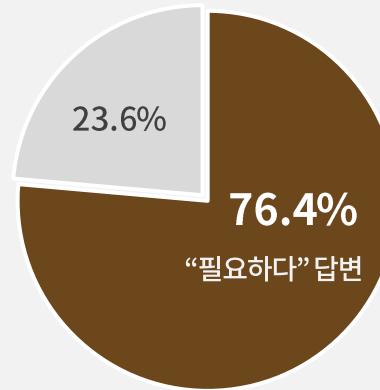


#### 원인

- 변호사 선임 비용의 부담
- 소액사건 소송의 간략한 절차로 인한 변호사 선임 필요성 X

2020년 기준, 민사소송 중 ‘나홀로 소송’ 72.6%

### 법률 서비스 IT 도입의 필요성 및 필요 이유 <중앙일보>



법률 서비스에 IT 도입이 필요하다고 답변, 응답자의 76.4%



과거보다 법률 정보에 접근방식이 용이해져 많은 사람들이 홀로 소송에 나서는 것으로 보이지만,  
합리적 증거를 제시하지 못하거나 법률 지식이 없어 **당사자의 권익을 제대로 구현하지 못하는 경우가 대다수이다.**

# 1. 문제 정의 및 연구 배경

## 1) 문제 정의

“전세사기, 깡통 전세, 주택 임대차 분쟁 증가...” <한국경제>

기사 中 일부 발췌 [1]

“ 저금리 기조가 오래 지속되면서 부동산 가격이 상승하고, 이로 인한 분쟁 또한 증가했다.  
부동산 분쟁은 민사 소송 중 약 30%를 차지할 정도로 그 빈도수가 높으며 분쟁의 대상이 될 뿐 아니라 가압류, 가처분 등 보전 처분은 대부분이 부동산을 대상으로 이뤄지는 것이 일반적이다. ”

“골치 앓이 부동산 소송, 경험 있는 조력자 도움 받아야...” <뉴스웍스>

기사 中 일부 발췌 [2]

“ 부동산 임대업, 임대차 계약자, 부동산 매매 등 부동산은 일상 생활에서 배제할 수 없는 요소 중 하나다. 때문에 부동산과 관련한 법정 분쟁이나 소송이 발생할 경우 어려운 부동산 용어 때문에 법적인 대응을 하기 어려운 부분들이 많이 있다. ”

부동산 분쟁은 일상생활과 밀접한 관련이 있고, 민사 소송 중 30%를 차지할 정도로 빈도수가 높다.

이는 부동산 관련 법률 정보 서비스 제공의 필요성을 시사한다.

# 1. 문제 정의 및 연구 배경

## 2) 연구배경

### 선행 연구

AI시대 LegalTech와 미래 법률시장[3]	<ul style="list-style-type: none"><li>• AI 도입으로 법률 서비스의 질, 국민 접근성 향상</li><li>• 법률 분야에서의 인공지능 활용도는 다른 분야에 비해 낮음</li></ul>
CNN, RNN을 이용한 법률 문서 분류[4] [5] [6]	<ul style="list-style-type: none"><li>• CNN, RNN을 이용한 법률 문서 분류 자동화</li><li>• 비정형화 문서, 일상적인 언어 사용시 정확도가 낮다는 한계점을 지님</li></ul>
주제어 태그를 이용한 판례 분류[7] 및 생활용어와 법률 용어간 대응관계[8]	<ul style="list-style-type: none"><li>• LDA 토픽 모델링 기반 주제어 태그</li><li>• 법률 용어가 낯선 일반인도 사용 가능</li><li>• 블로그의 태그 정보를 활용하여 검색하고자 하는 생활용어에 대응하는 법률용어 탐색</li></ul>

법률 분야에서 다양한 연구들이 진행되고 있으며,  
판례를 활용하여 일반인에게 법률 정보를 제공하는 정보시스템의 필요성 대두

### 기존 서비스



법제처  
국가법령정보센터

LegalSearch

LBox

CaseNote

#### 국내 4개 주요 법률 정보 검색 서비스

- 사건번호나 전문 법률 용어를 질의어로 입력해야만 검색 가능
- 문장이 아닌, 키워드 입력 형태로 검색
- 검색 결과 내에서도 본인의 사건 내용에 맞는 판례를 찾기 위해 전문가 수준의 법률 지식이 요구되므로 일반인의 활용도가 떨어짐

일반인에게 진입장벽이 높은 기존 정보 검색 서비스를 보완하기 위한,

법률용어가 아닌 **일반용어로 구성된 문장으로도 법률 정보 검색이 가능한 시스템**의 필요성 확인

## 2. 연구 목표

### 1) 세부 목표 및 시스템 흐름도

#### 세부 목표

##### 법률 용어 자동 변환

법률 용어가 아닌, 일반인들이 사용하는 언어로 검색해도 의미 분석 가능

##### 유사도 분석

기존 알고리즘을 활용한 법률 문서 간 유사도 분석을 통해 유사 판례 및 참조 가능한 법 조항 추출

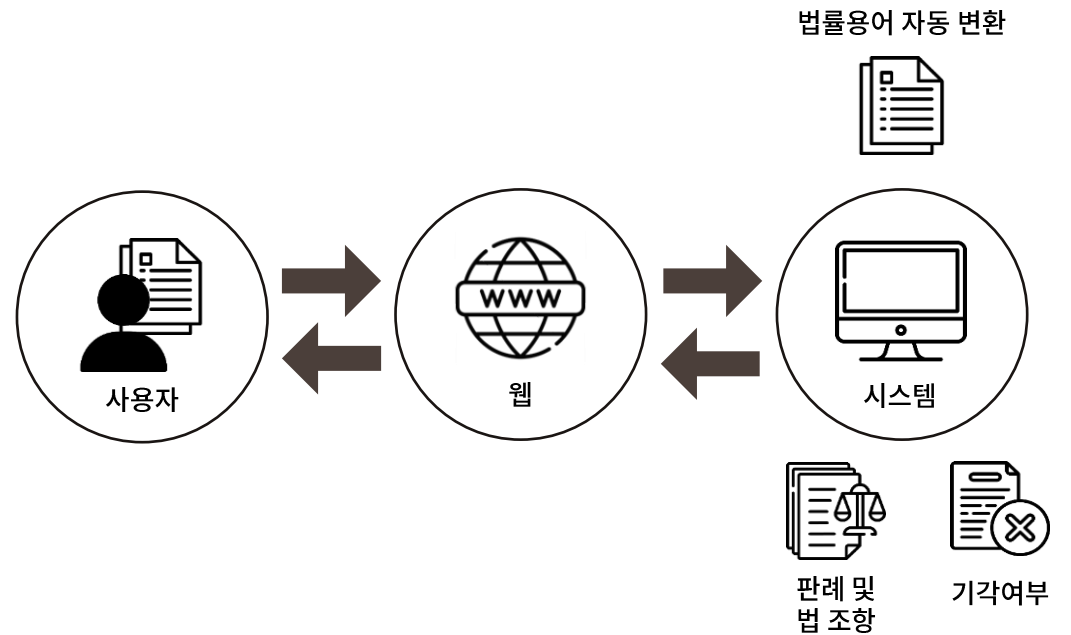
##### 기각 여부 판단

Input과 유사한 상위 10개 판례들의 판결문을 통한 기각 여부 예측

##### 웹 서버 구축

사용자가 접근하기 쉽고 이용하기 쉬운 웹 구축

#### FLOW CHART



법률 취약계층인 일반인을 대상으로 사건 내용과 관련된  
판례, 참조 가능한 법 조항, 기각 여부 예측



법률 취약 계층의 재판 정보 비대칭 문제 완화

### 3. 연구 과정

1) 연구 방법론

#### 1. 데이터 수집

국가법령정보센터

LAWnB

Beautiful Soup를 이용하여 XML 형태의 하급심 판례 크롤링  
하급심 부동산 관련 손해배상 판례 340개

Selenium, Beautiful Soup를 이용하여 지방법원 판례 동적 크롤링  
지방 법원 부동산 관련 손해배상 판례 668개

#### 2. 데이터 전처리

- ✓ 데이터의 구조가 모두 동일하지 않기 때문에 유사도 비교 분석 기준인 판시사항이 NaN값인 행 삭제
- ✓ 불용어 리스트 구축 (최빈 단어들 중 내용상 무의미한 723개 단어) → 불용어 제거 및 명사 추출

ID	판시사항	판결요지	참조조문	참조판례	주문	청구취지	판시+주문
0 165171	甲이 배우자 乙의 성기능 장애로 인하여 사실혼관계가 단기간에 파탄되었다는 이유로 乙...	甲이 배우자 乙의 성기능 장애로 인하여 사실혼관계가 단기간에 파탄되었다는 이유로 乙...	민법 제806조	1. 피고(반소원고) 1은 원고(반소피고)에게, 가. 20,000,000원과 ...	1. 본소 가. 주위적 청구취지피고(반소원고, 이하 '피고'라 한다)들은 연대하여...	甲이 배우자 乙의 성기능 장애로 인하여 사실혼관계가 단기간에 파탄되었다는 이유로 乙...	

불용어		711	고려하면
702	점	712	관련
703	증	713	까닭
704	바	714	누가
705	운	715	달다
706	후	716	대해
707	갓다	717	더욱이
708	거리	718	도달
709	견지	719	돼다
710	결론	720	되다

불용어 리스트 中 발췌

### 3. 연구 과정

#### 1) 연구 방법론

### 3. 수집 데이터 벡터화

CountVectorizer

word2Vec, ELMo

✓ TF-IDF

단어가 문서 내 등장한 횟수만을 count → 의미 없는 단어에 높은 수치를 부여할 가능성 존재

국문으로 구성된 법률 관련 사전 학습 언어모델의 부재로 인해 성능을 기대할 수 없음

법률 용어로 이루어진 판례 데이터를 토대로 임베딩 벡터 생성 후,  
각 단어들마다 다른 가중치를 곱해서 표현

### 4. 벡터 간 유사도 계산

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|}$$

→ cosine similarity 값 기준 내림차순으로 유사 판례 10개 정렬



## 1) 연구 방법론

## PCA



## t-SNE

정보를 대부분 유지하면서 차원을 줄이는 통계적 방법으로  
선형 방식으로 정사영했기 때문에 직관적인 해석이 어려워 시각화로 부적합

t 분포를 활용해 고차원인 본래 공간에 존재하는 데이터의 이웃간 거리를 최대한 보존하며 저차원에서 학습하는 방법

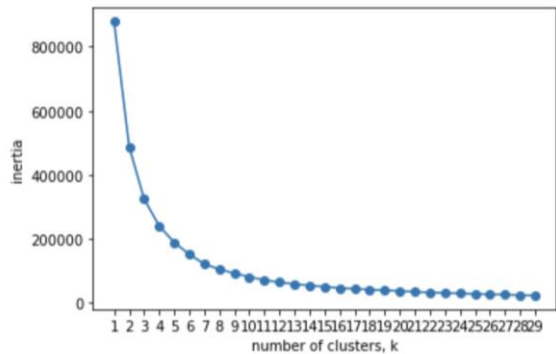
Wordcloud 내  
빈도수 높은 단어를  
기준으로  
일반 용어에서  
법률 용어로의 맵핑

법률 용어	일반 용어
채권자(131)	돈을 받아야하는 사람, 돈을 받을 권리가 있는 사람, 청구할 수 있는 권리를 가진 사람, 돈을 빌려준 사람, 이자를 받는 사람, 소송 당사자, 빚을 받을 권리가 있는 사람, 돈을 돌려받아야하는 사람
임대인(92)	집주인, 주인 아줌마, 주인 아저씨, 주인
임차인(90)	세입자, 세들어 사는 사람
...	...

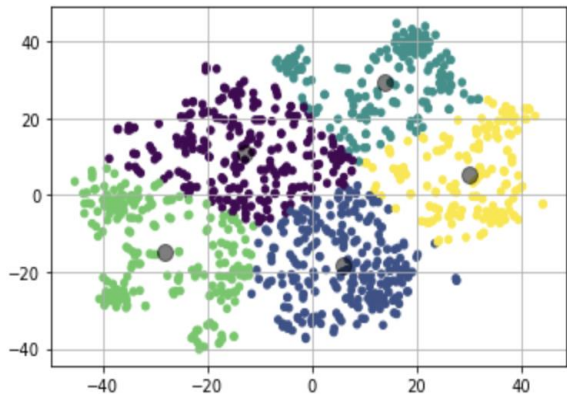
# 4. 연구 결과

## 1) 연구 결과

### Clustering 과정



Inertia 값에 따라 클러스터 개수 k=5 채택



K-means로 클러스터 5개 선정

num	clusters	cluster_num
0	0 [6145, 4100, 6196, 6205, 4158, 6206, 66, 4171,...	230
1	1 [8, 2065, 2072, 25, 31, 50, 4160, 6209, 6224, ...	247
2	2 [7, 6155, 2060, 6170, 2087, 6207, 88, 4187, 21...	168
3	3 [4116, 6213, 6265, 162, 6331, 4300, 438, 6607,...	207
4	4 [4133, 2120, 2143, 4221, 2180, 6286, 144, 4253...	156

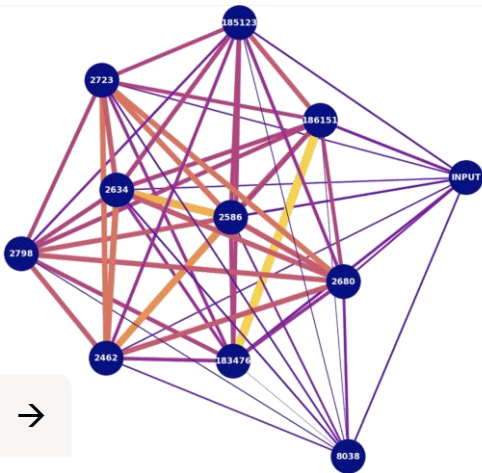
각 클러스터의 판례ID, 군집내 Data개수

### 네트워크 시각화 과정

	INPUT	186151	183476	2586	2680	185123	8038	2462	2723	2798	2634
INPUT	1.000000	0.393234	0.386296	0.345622	0.324408	0.320588	0.315352	0.283420	0.282865	0.282107	0.277698
186151	0.393234	1.000000	0.882939	0.564848	0.480554	0.566155	0.219787	0.521017	0.483453	0.521458	0.525257
183476	0.386296	0.882939	1.000000	0.494503	0.452975	0.506115	0.191215	0.458648	0.432201	0.514353	0.425626
2586	0.345622	0.564848	0.494503	1.000000	0.615081	0.528015	0.330645	0.738306	0.618501	0.574150	0.803148
2680	0.324408	0.480554	0.452975	0.615081	1.000000	0.443222	0.382129	0.606361	0.618745	0.590953	0.598789
185123	0.320588	0.566155	0.506115	0.528015	0.443222	1.000000	0.235319	0.440780	0.490486	0.365255	0.535631
8038	0.315352	0.219787	0.191215	0.330645	0.382129	0.235319	1.000000	0.291962	0.310825	0.248656	0.318995
2462	0.283420	0.521017	0.458648	0.738306	0.606361	0.440780	0.291962	1.000000	0.619308	0.572113	0.681075
2723	0.282865	0.483453	0.432201	0.618501	0.618745	0.490486	0.310825	0.619308	1.000000	0.500594	0.602472
2798	0.282107	0.521458	0.514353	0.574150	0.590953	0.365255	0.248656	0.572113	0.500594	1.000000	0.520769
2634	0.277698	0.525257	0.425626	0.803148	0.598789	0.535631	0.318995	0.681075	0.602472	0.520769	1.000000

← 상위 유사 판례 10개와 input의 인접행렬 생성

생성한 인접행렬을 기반으로 노드간 네트워크 시각화 →



# 4. 연구 결과

## 2) Prototype

Dash

127.0.0.1:8050

업데이트

분야 별 판례 조회

사건 내용에 따른 법률 정보 검색

분야 선택

결과 보기

관심 있는 분야를 선택하세요.

☒ 공사, 철거, 분양 ( 230 )


☐ 매매, 보증금, 배상 ( 247 )

☐ 주택, 계약, 신축 ( 168 )

☐ 인감, 명의, 근저당 ( 207 )

☐ 판결, 권리금, 소송 ( 156 )

선택하신 ‘공사, 철거, 분양’ 분야 내 자주 등장하는 단어



선택하신 ‘공사, 철거, 분양’ 분야 내 판례 230개

1) 경작지인도 및 손해배상 청구 사건  
[https://www.law.go.kr/precInfoP.do?mode=0&precSeq=71848&vSct=\\*](https://www.law.go.kr/precInfoP.do?mode=0&precSeq=71848&vSct=*)

2) 손해배상 및 토지 인도 청구 사건  
[https://www.law.go.kr/precInfoP.do?mode=0&precSeq=71871&vSct=\\*](https://www.law.go.kr/precInfoP.do?mode=0&precSeq=71871&vSct=*)

3) 토지소유권 이전 등기말소(본소), 손해배상(반소) 청구 사건  
[https://www.law.go.kr/precInfoP.do?mode=0&precSeq=73539&vSct=\\*](https://www.law.go.kr/precInfoP.do?mode=0&precSeq=73539&vSct=*)

더보기..

# 4. 연구 결과

## 2) Prototype

Dash

127.0.0.1:8050

업데이트

분야 별 판례 조회

사건 내용에 따른 법률 정보 검색

현재 처한 상황

결과 보기

사건 내용

참조 가능 법조항 및 예상 기각 여부

상위 유사 판례

유사 판례 네트워크

법률 용어 자동 변환

확인

저는 상가 빌딩 내에서 음식점을 운영하고 있습니다. 임대차계약이 끝나기 전에 다른 사람이 제가 운영하는 음식점을 임차하기를 원해서 권리금을 받기로 하였지만 빌딩 주인이 그 과정에서 그 사람에게 여러 서류 제출을 요구하고 원래 세입자가 내는 돈보다 많이 돈을 내라고 하여 타협이 취소되었습니다. 저는 빌딩 주인이 권리금 돌려받는 것을 방해했다고 생각하여 손해배상을 청구하려 합니다.

참조 가능 법 조항

1) 상가건물 임대차 보호법 제 2조  
<https://www.law.go.kr/lsLinkProc.do?lsClsCd=L&lsNm=%EC%83%81%EA%B0%80%EA%B1%B4%EB%AC%BC%20%EC%9E%84%EB%8C%80%EC%B0%A8%EB%B3%B4%ED%98%B8%EB%B2%95&lsId=prec20171026&joNo=000200&efYd=20171026&mode=11&lnkJoNo=undefined>

2) 상가건물 임대차 보호법 제 8조  
<https://www.law.go.kr/lsLinkProc.do?lsClsCd=L&lsNm=%EC%83%81%EA%B0%80%EA%B1%B4%EB%AC%BC%20%EC%9E%84%EB%8C%80%EC%B0%A8%EB%B3%B4%ED%98%B8%EB%B2%95&lsId=prec20171026&joNo=001000&efYd=20171026&mode=11&lnkJoNo=undefined>

더보기..

예상 기각 여부

유사 판례 10개 中 6개 기각, 3개 일부 승소, 1개 승소

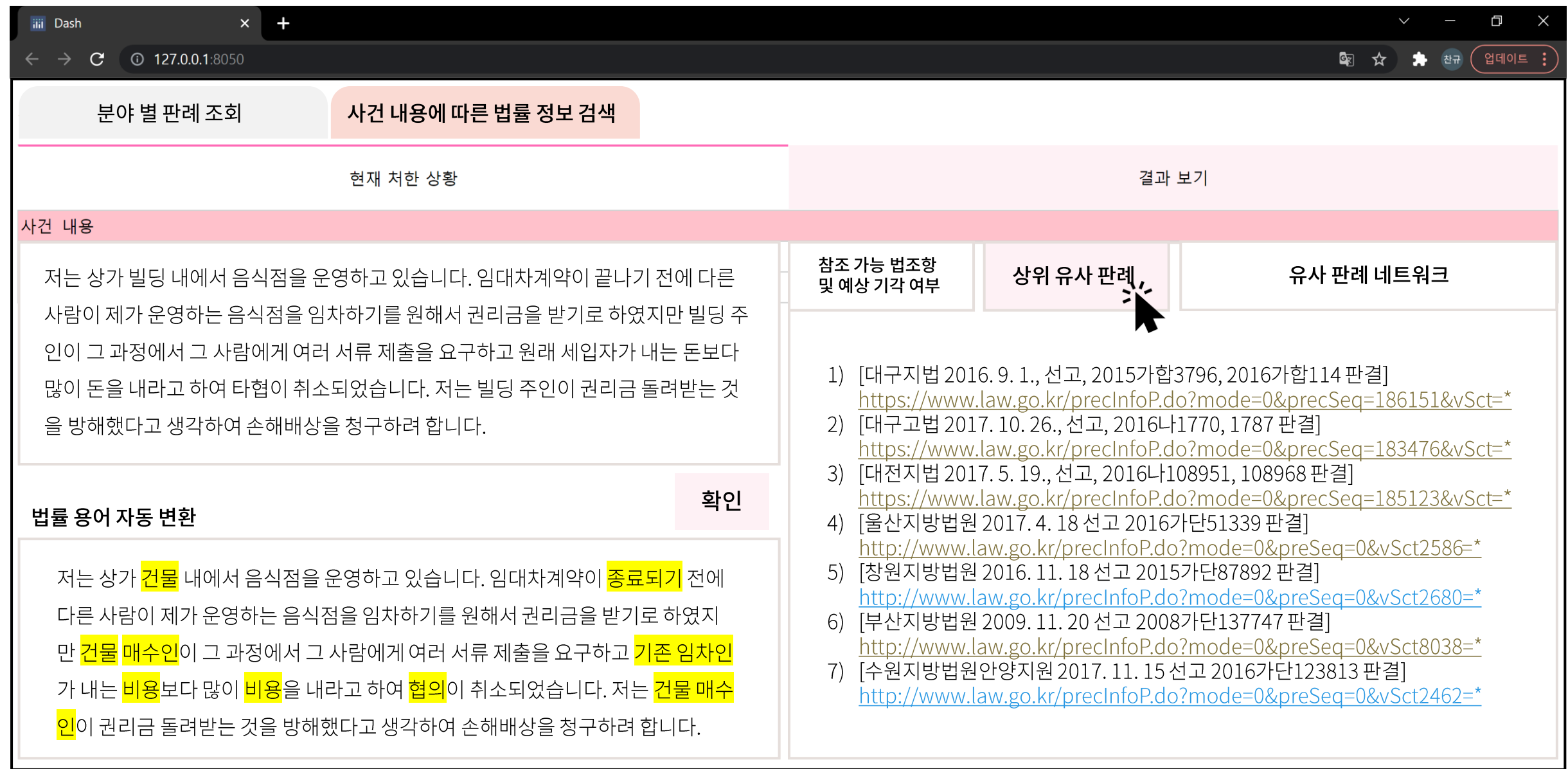
→ 본 사건은 ‘기각’ 될 것으로 예측됩니다.

법률 용어 자동 변환

저는 상가 건물 내에서 음식점을 운영하고 있습니다. 임대차계약이 종료되기 전에 다른 사람이 제가 운영하는 음식점을 임차하기를 원해서 권리금을 받기로 하였지만 건물 매수인이 그 과정에서 그 사람에게 여러 서류 제출을 요구하고 기존 임차인이 내는 비용보다 많이 비용을 내라고 하여 협의이 취소되었습니다. 저는 건물 매수인이 권리금 돌려받는 것을 방해했다고 생각하여 손해배상을 청구하려 합니다.

# 4. 연구 결과

## 2) Prototype



# 4. 연구 결과

## 2) Prototype

Dash

127.0.0.1:8050

업데이트

분야 별 판례 조회

사건 내용에 따른 법률 정보 검색

현재 처한 상황

결과 보기

사건 내용

참조 가능 법조항 및 예상 기각 여부

상위 유사 판례

유사 판례 네트워크

법률 용어 자동 변환

확인

저는 상가 빌딩 내에서 음식점을 운영하고 있습니다. 임대차계약이 끝나기 전에 다른 사람이 제가 운영하는 음식점을 임차하기를 원해서 권리금을 받기로 하였지만 빌딩 주인이 그 과정에서 그 사람에게 여러 서류 제출을 요구하고 원래 세입자가 내는 돈보다 많이 돈을 내라고 하여 타협이 취소되었습니다. 저는 빌딩 주인이 권리금 돌려받는 것을 방해했다고 생각하여 손해배상을 청구하려 합니다.

저는 상가 건물 내에서 음식점을 운영하고 있습니다. 임대차계약이 종료되기 전에 다른 사람이 제가 운영하는 음식점을 임차하기를 원해서 권리금을 받기로 하였지만 건물 매수인이 그 과정에서 그 사람에게 여러 서류 제출을 요구하고 기존 임차인이 내는 비용보다 많이 비용을 내라고 하여 협의이 취소되었습니다. 저는 건물 매수인이 권리금 돌려받는 것을 방해했다고 생각하여 손해배상을 청구하려 합니다.

판례간 유사도 大

판례간 유사도 小

## 5. 토의, 한계 및 미래 연구

### 1) 토의, 한계 및 미래 연구

#### › 토의

##### 연구의 타당성

문서 분석에 적합한  
알고리즘인 TF-IDF[5]를 활용하고,  
PIP Loss[9]를 통해 과적합을 방지함

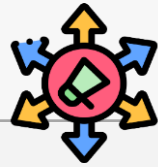
##### 연구의 유의성

일반인 사용자가 필요로 하는 관련 법률 정보  
접근을 용이하게 하는 시스템을 구축함으로써  
정보 비대칭 문제 완화

##### 연구의 기대효과

- ① 법률 정보가 필요한 일반인들에게  
정보 전달의 시간/비용적 절감을 기대할 수 있음
- ② 법조인들의 참조 조항, 판례 검색 등  
자료 조사에서의 활용성을 기대할 수 있음

#### › 한계 및 미래 연구



연구 범위의 확장

- 민사소송을 다루며 그 중에서도 손해배상 분야를 채택
  - 손해배상 분야 내 부동산 관련 소송만을 다룸
- 연구 과정을 표준화하여 다양한 소송 분야에도 적용이 가능하도록  
**연구 범위를 확장해볼 가치가 충분함**



데이터 수집의 한계

- 공개된 판례 데이터가 전체의 0.6%, 그 안에서 부동산 관련 건 10%이하
  - 연구의 정확도를 위해 다량의 데이터 필요
- 사법부의 판례 데이터 전체 공개를 통해 **다양한 각도에서의 법률  
인공지능의 발전** 가능

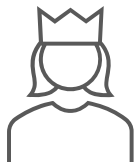


손해배상액  
예측의 어려움

- 손해배상액을 도출해내는 방법 또는 공식의 부재
  - 원고,피고의 사회적 지위나 상황에 따라 금액을 산정하는 기준이 다름
  - 판사의 주관적인 견해를 완전히 배제할 수 없음
- 소송을 준비하는 **사용자들이 가장 궁금해하는 부분**이라고 판단하여 이  
를 더 연구 해볼 가치가 충분함

# 6. 역할 분담

## 1) 역할 분담



지승빈 (팀장)

- 프로젝트 총괄
- 중간, 최종 발표



박병현

- 데이터 전처리
- 알고리즘 구현



이찬규

- 데이터 크롤링
- 개발환경 구축



임채균

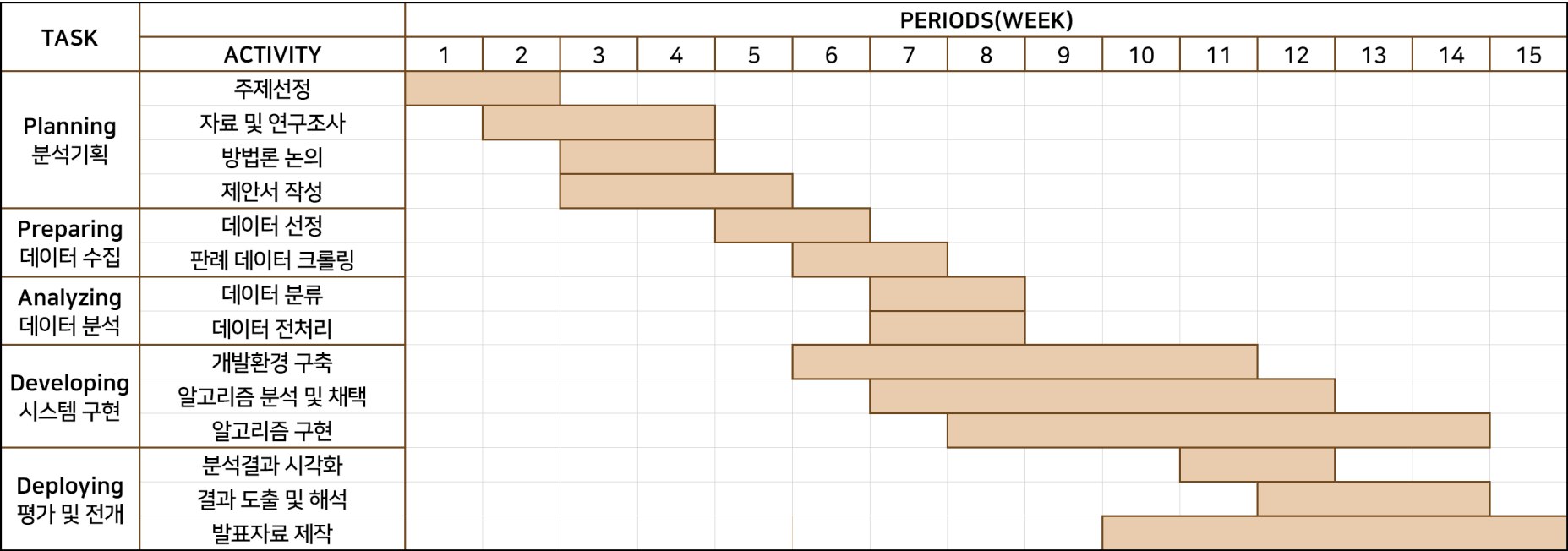
- 선행연구 및 자료조사
- 데이터 수집 및 관리



지유하

- 선행연구 및 자료조사
- 발표자료 ppt 작성

Gantt Chart





## 7. References

### 1) 참고 문헌

---

- [1] 김준우, ‘소송으로 원하는 결과 얻기 위해서는 가압류, 가처분 신청이 중요’ (2021)
- [2] 장원수, ‘골치 앓이 부동산 소송, 경험있는 조력자 도움 받아야’ (2018)
- [3] 김승래 (2020). AI시대 리걸테크의 발전과 미래 법률시장의 변화 모색. 법이론실무연구, 8(3),85-128.
- [4] 이지훈, 배득현, 신호준, 김재윤, 이혁준 (2018). CNN을 사용한 법률문서 분류 시스템에 관한 연구. 한국통신학회 학술대회논문집, 939-940
- [5] 이지훈, 엄태현, 이혁준 (2019). TF-IDF와 CNN을 사용한 법률문서 분류 시스템에 관한 연구. 한국통신학회 학술대회논문집, 982- 983
- [6] 이지훈, 이혁준 (2019). RNN을 사용한 법률문서 분류 시스템에 관한 연구. 한국통신학회 학술대회논문집, 1027- 1028
- [7] 심준식, 김형중 (2017). LDA 토픽 모델링을 활용한 판례 검색 및 분류 방법. 전자공학회논문지, 54(9), 67-75
- [8] 김지현, 이종서, 이명진, 김우주, 홍준석.(2012).법령정보 검색을 위한 생활용어와 법률용어 간의 대응관계 탐색 방법론.지능정보연구,18(3),137-152.
- [9] Zi Yin and Yuanyuan Shen. 2018. On the dimensionality of word embedding. In Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18). Curran Associates Inc., Red Hook, NY, USA, 895–906.

---

THANK YOU 😊

---