LEGALSEARCHLM: Rethinking Legal Case Retrieval as Legal Elements Generation

Chaeeun Kim[1*]   **Jinu Lee**[2]   **Wonseok Hwang**[1,3]

[1]LBOX   [2]University of Illinois Urbana-Champaign   [3]University of Seoul

Abstract

Legal Case Retrieval (LCR), which retrieves relevant cases from a query case, is a fundamental task for legal professionals in research and decision-making. However, existing studies on LCR face two major limitations. First, they are evaluated on relatively small-scale retrieval corpora (e.g., 100-55K cases) and use a narrow range of criminal query types, which cannot sufficiently reflect the complexity of real-world legal retrieval scenarios. Second, their reliance on embedding-based or lexical matching methods often results in limited representations and legally irrelevant matches. To address these issues, we present: (1) **LEGAR BENCH**, the first large-scale Korean LCR benchmark, covering 411 diverse crime types in queries over 1.2M candidate cases; and (2) **LEGALSEARCHLM**, a retrieval model that performs legal element reasoning over the query case and directly generates content containing those elements, grounded in the target cases through constrained decoding. Experimental results show that LEGALSEARCHLM outperforms baselines by 6 – 20% on LEGAR BENCH, achieving state-of-the-art performance. It also demonstrates strong generalization to out-of-domain cases, outperforming naive generative models trained on in-domain data by 15%.

1 Introduction

Legal AI has increasingly gained attention from legal professionals to raise the productivity of their work. Among various legal applications, *Legal Case Retrieval* (LCR) (Feng et al., 2024; Deng et al., 2024b; Su et al., 2024; Deng et al., 2024a; Li et al., 2023a; Xiao et al., 2021; Li et al., 2023b,c; Zhang et al., 2023), which identifies relevant precedents for a given case, plays a particularly crucial role in maintaining judicial fairness and supporting the decision-making process of legal experts.

However, previous work on the legal case retrieval task has clear limitations. (1) Most evaluations have been conducted in small-scale settings, where a predefined set of candidate documents is provided for each query and the queries do not sufficiently reflect the diverse criminal case types (Ma et al., 2021; Li et al., 2023b,c; Xiao et al., 2018; Li et al., 2023d). (2) The tested retrieval models are limited to embedding-based similarity and lexical matching approaches, with the former struggling to capture the rich semantics of legal literature as it compresses complex documents into fixed-size vectors, and the latter often leading to unfocused matches due to a lack of semantic understanding (Wang et al., 2023; Magesh et al., 2024; Kim et al., 2024).

To address the first, we present LEGAR BENCH (**Lega**l Case **R**etrieval **Bench**mark), the first large-scale Korean LCR benchmark, comprising two dataset versions tailored to different evaluation needs: (1) LEGAR BENCH*Standard* is designed for comprehensive assessment across a broad range of crime categories. The queries cover 411 distinct crime types and are evaluated over a retrieval pool of 1.2M cases. To achieve this, we systematically construct the benchmark using crime types based on statutory provisions, even further detailed than official charge titles used in courts (e.g., defamation by fact disclosure – Article 307(1), or false allegation – Article 307(2), both sharing the charge title *defamation*) (Section 2.1). (2) LEGAR BENCH*Stricter* evaluates stricter relevance criteria than LEGAR BENCH*Standard*. It considers more factual details and legal issues within the same crime type that can affect the final judgment or sentence, which is crucial for legal practitioners. To enable this, we annotate the pool of 170K cases across 160 crime types using 102 crime-specific legal factors and 443 corresponding options (Section 2.2).

To address the second, we shift our focus on the

---

*Correspondence to: chaechaek1214@gmail.com

	Language	Crime types of query	Query case	Retrieval pool	Target case per query
COLIEE2024	English	-	400	1,734 (per query)	-
LeCaRD	Chinese	-	107	100 (per query)	10.33
LeCaRDv2	Chinese	-	800	55,192 (per query)	20.89
LEGAR BENCH$_{Standard}$	Korean	411	411*N	1,226,814	200
LEGAR BENCH$_{Stricter}$	Korean	160	~15,777	169,230	14.69

Table 1: Scale comparison of LCR benchmarks. LEGAR BENCH expands the retrieval pool and query coverage. N indicates query set expandability via similar case groups.

LCR task from sequence-to-sequence matching to generating the key legal elements that determine relevance by proposing LEGALSEARCHLM. It directly generates important legal elements via constrained decoding and returns the documents that contain them (Li et al., 2024; Kim et al., 2024; Li et al., 2023e,f; Bevilacqua et al., 2022), mitigating the shortcomings of traditional retrievers like BM25 and dense embeddings. Given the specificity of the legal domain, we adopt diverse strategies including first token-aware generation and self-supervised fine-tuning (SSFT) (Section 3).

We evaluate LEGALSEARCHLM on LEGAR BENCH, comparing it against strong baselines including lexical matching and embedding-based methods from both general-purpose and legal-domain models. Our experimental results show that LEGALSEARCHLM outperforms the best baseline by 6% in precision. Our training strategy leads to significantly better generalization, achieving a 15% improvement over generative retrieval models trained with naive identifiers on in-domain data (Section 5).

In summary, our contributions are as follows.

- We introduce the first Korean LCR benchmark, LEGAR BENCH, which has the largest and most diverse criminal cases.

- We present LEGALSEARCHLM that generates legal elements that should be included in target case via first-token-aware decoding.

- Our LEGALSEARCHLM achieves state-of-the-art performance on LEGAR BENCH, demonstrating remarkable *generalization ability* on unseen crime types.

2   LEGAR BENCH

In this section, we describe relevance criteria and the construction process of LEGAR BENCH.

LEGAR BENCH features the most comprehensive set of query and target cases (See Table 1),

Categories	#Crime types of queries	
	Standard	Stricter
Traffic offenses	13	9
Fraud	21	8
Injury or Violence	31	19
Sexual crime	132	111
Defamation or Insult	8	6
Finance or Insurance	5	1
Drug	5	4
Murder	2	2
Theft or Robbery	38	-
Obstruction of Business	13	-
Destruction	5	-
Threat	11	-
Criminal trespass	15	-
Embezzlement or Breach of trust	15	-
Gambling	7	-
Negligent homicide and injury	6	-
Obstruction of right	5	-
Child abuse or School violence	10	-
Medical or Food drug	11	-
Corporation	3	-
Bribery	3	-
Car	2	-
Labor or Employment	11	-
Industrial or Serious accidents	4	-
Military duty or law	2	-
Consumer or Fair trade	1	-
Arrest or Detention	1	-
Intellectual property	3	-
IT or Privacy	2	-
Misdemeanor	1	-
Sexual norms	1	-
Tax, Administ, Const law	14	-
Other criminal offenses	10	-
Total	**411**	**160**

Table 2: Statistics of Crime Typology and Corresponding Criminal Types. LEGAR BENCH$_{Standard}$ includes 411 query types across 33 crime categories, while LEGAR BENCH$_{Stricter}$ covers 160 query types across 8 categories.

with relevance criteria rigorously defined by the lawyers involved in the construction process. It offers two dataset versions based on different evaluation needs, LEGAR BENCH$_{Standard}$ and LEGAR BENCH$_{Stricter}$.

2.1   LEGAR BENCH$_{Standard}$

LEGAR BENCH$_{Standard}$ is designed to provide a comprehensive assessment of retrieving diverse crime categories, and the enhancement of overall
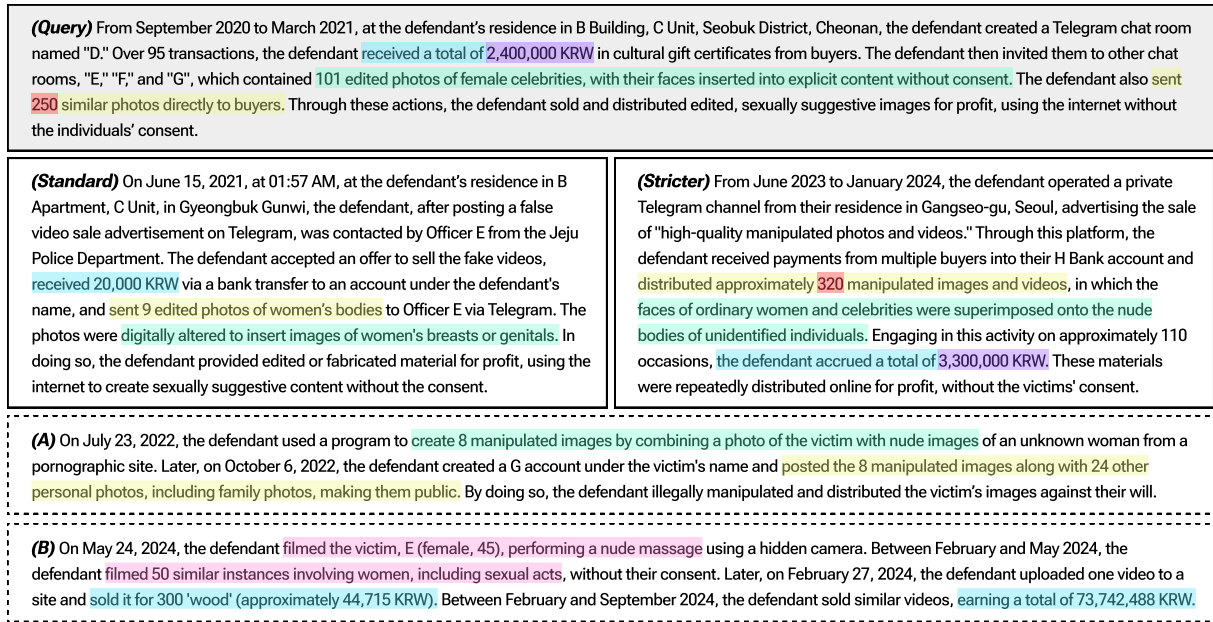
Figure 1: Examples of Relevance Cases. **(Query)** is a query case on distributing false images/videos for profit. The Blue Highlight indicates profit, the Yellow Highlight represents the creation of false images/videos, and the Green Highlight denotes distribution—the three key legal elements of the crime. Both **(Standard)** and **(Stricter)** satisfy the three elements, and **(Stricter)** additionally meets the requirements concerning the scale of distributed images/videos (Red Highlight) and the total financial gains obtained (Purple Highlight). **(A)** and **(B)** are not target cases, as **(A)** distributed a false image without intending to obtain financial gains, and **(B)** committed the offense for financial gain through the unlawful filming of real footage, not the creation of false images.
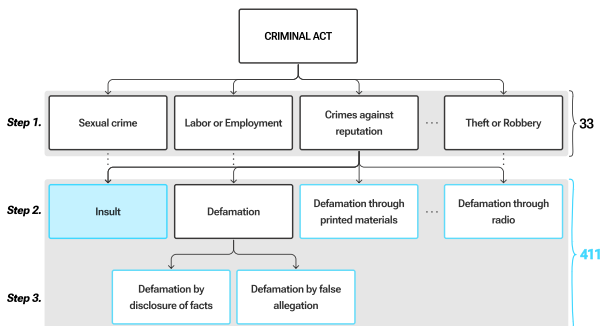

Figure 2: Examples of the construction process for each step in LEGAR BENCH*Standard*. **Step 1** defines major crime categories based on Korean Criminal Act. **Step 2** refines these categories using charge titles, and **Step 3** further specifies them based on statutory provisions.

system performance.

2.1.1 Definition of Standard Relevance

We define standard relevance based on the charge title and statutory provision, where the former refers to the formal name of the offense, and the latter indicates the specific statutory article applicable to that charge. For example, as shown in Figure 1 on sexual crime, for the query case **(Query)** on dis-

tributing false sexual images/videos for profit, the standard target case **(Standard)** satisfies the three statutory elements: 1. creation of false sexual images/videos, 2. intent for profit, and 3. distribution. Cases like **(A)**, which are not for profit, or **(B)**, which concern illegal sexual video filming rather than false sexual image creation, cannot be considered target cases, since they are distinct crimes governed by different laws.

2.1.2 Data Construction

To ensure comprehensive coverage of the diverse and complex legal literature, we employ a top-down approach, systematically categorizing crimes based on Korean Criminal Act.

Step 1: Construction of Crime Typology. We establish a crime typology to categorize various types of crimes in criminal cases. As shown in Step 1 of Figure 2, we define major categories based on the structure of the Korean Criminal Act, such as sexual crimes, labor or employment offenses, crimes against reputation, and theft or robbery. Appendix 5 lists a total of 33 crime categories.
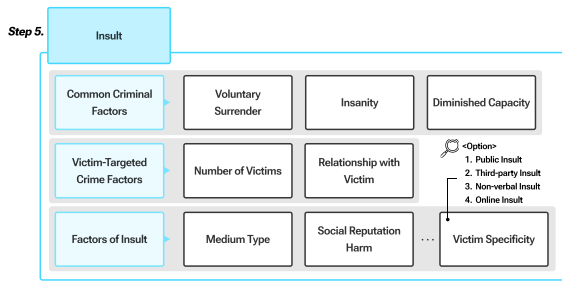
Figure 3: Examples of the construction process of LEGAR BENCH*Stricter*. Each crime type (e.g., Insult) includes specifically defined factors (in boxes filled with sky blue) and sub-factors (in boxes outlined in black). Cases are annotated by mapping all sub-factors to corresponding predefined options.

Step 2: Assignment of Charge Titles. We construct the set of crime charge titles that can occur within each crime category. A charge title is the official name used in legal documents, such as indictments or complaints, to describe a specific offense. As shown in Figure 2, crimes against reputation can be expanded into related charge titles (sub-categories), such as defamation, defamation through printed materials, defamation through radio, insults, etc. While charge titles are determined by statutory provisions that apply, there is no one-to-one correspondence between these two because some charge titles correspond to multiple provisions with subtle differences. This ambiguity is resolved by further refinement in the next step.

Step 3: Refinement from Statutory Provisions. To better reflect the subtle difference of the body of the crime between provisions grouped within the same charge title, we refine charge titles to the level of individual statutory provisions when such refinement is both possible and meaningful. Figure 2 shows how defamation can be divided according to distinct laws, such as defamation by disclosure of facts and defamation by allegation of false facts. Finally, the standard similar groups are formed by combining the results of Step 2 and Step 3, as shown in the skyblue-bordered box of Figure 2. *As a result, LEGAR BENCH*Standard *contains 411 similar groups across 33 categories.*

Step 4: Case Mapping. We automatically process 1.2M (1,226,814) criminal cases, mapping them to their respective groups based on the charge

title and statutory provisions annotated for each group. This process successfully maps 1M cases (1,052,506), which account for 85.79% of our total criminal cases, *enabling evaluation on the majority of criminal cases through our LEGAR BENCH*Standard.

2.2 LEGAR BENCH*Stricter*

2.2.1 Definition of Stricter Relevance

For stricter case similarity, we expand the scope from facts to include claims, reasoning, sentencing factors, and conclusions sections from the case, aiming to provide a more comprehensive view of the process of judges. Stricter relevance further requires *factual details* that do not affect the type of charge, but might affect the final judgment or the sentence. Examples include the severity of the crime, the relationship between the defendant and the victim, situational information, and arguments made by defendants. For instance, while making only a few fake images and selling them for 20 dollars is ruled under the same crime with making hundreds of fake videos with thousands of dollars of profit, the stricter factual relevance between the two cases is low (See the red and purple highlights in Figure 1). Also, if two assault defendants make the same claim of self-defense but only one of them is accepted by the judge, these two cases should also be distinguished. Five legal experts have annotated these important factors that determine the stricter relevance between cases (Appendix C.1).

2.3 Dataset Construction

Step 5: Define Detailed Factors. We construct LEGAR BENCH*Stricter* starting from 160 similar groups across 8 crime categories in LEGAR BENCH*Standard*. First, we define sets of factors to be further considered for each group in the standard set. Figure 3 shows that the standard group "Insult" is associated with Common Criminal Factors, Victim-Targeted Crime Factors, and Factors of Insult. Next, we identify detailed sub-factors and create options for each of them. Finally, based on the defined factors, sub-factors, and options for each standard group, we annotate the cases belonging to each standard group using GPT-4o. The full list of sub-factors is shown in Table 7.

Step 6: Case Grouping. As a result of Step 5, we obtain (sub-factor, option) pairs for each case across all sub-factors required for each standard group. The following grouping algorithm is then

applied to find cases with the highest factual relevance. We created one stricter group for each standard group, resulting in 160 queries (See Table 1 for details).

Algorithm 1 Stricter Relevance Group

1: **Input:** case_data, subfactor-option pair_list
2: **Output:** grouped_cases
3: **for** each case in case_data **do**
4:     key = generate_key(subfactor-option pair_list)
5:     group[key].append(case)
6: **end for**
7: **if** any group has 2 or more cases **then**
8:     return the group
9: **end if**
10: **for** r = number of subfactors to 1 **do**
11:     **for** each case in case_data **do**
12:         key = generate_key(subfactor-option pair_list[:r])
13:         group[key].append(case)
14:     **end for**
15:     **if** any group has 2 or more cases **then**
16:         return the group
17:     **end if**
18: **end for**
19: **return** None

---

3   LEGALSEARCHLM

Previous approaches to LCR rely on either embedding-based or lexical matching methods (Magesh et al., 2024). However, embedding models often lose important details by compressing lengthy legal texts into single vectors, while lexical methods struggle to distinguish legally important information from noise. To better reflect how legal experts assess relevance—focusing on specific legal elements that constitute the crime, rather than subtle details like dates, geographic names, or place names—there is a need for a more sophisticated retrieval approach.

In this work, we introduce LEGALSEARCHLM, which addresses the pitfalls of sequence-to-sequence matching by adopting the generative retrieval paradigm (Bevilacqua et al., 2022; Li et al., 2023f,e; Kim et al., 2024). Given a query case, LEGALSEARCHLM generates relevant *legal elements* as keys for the target documents. By modeling retrieval as language modeling, it effectively mitigates the fixed-dimension embeddings' information loss problem and the lexical match's lack

of deep semantic understanding, which are both crucial in LCR.

3.1   Training

Extracting legal elements.   We define legal elements are atomic facts that can influence the final judgment. As proven effective by Min et al. (2023); Chen et al. (2024a); Cai et al. (2024), we extract legally valid elements by prompting an LLM.

However, naively training the model using extracted legal elements is insufficient for successful retrieval. This is especially true at inference time, where the decoding process is constrained by a predefined FM-index that forces the model to generate sequences appearing exactly in the corpus. In this framework, early decoding decisions are highly critical; for instance, generating "dates" or "locations" of the crime as the first token may unintentionally steer the decoding process toward irrelevant paths. This has the unintended effect of filtering based on information unrelated to legal relevance, thereby discarding documents that may contain key legal elements while overemphasizing those that happen to include the specific date or location.

To address this problem, we construct synthetic examples that begin with legally informative tokens from legal elements using few-shot LLMs. Figure 4 illustrates the generation of first-token-aware legal elements at inference time.

Data collection.   As LEGALSEARCHLM formulates LCR as legal elements generation, it needs finetuning to generate appropriate legal elements given the query case. However, as relevance annotation in the legal domain is costly, training the model on a large set of query-target pairs is not feasible.

As a solution, we construct a training dataset in a self-supervised manner (Lewis et al., 2020). Specifically, we use query case as inputs and legal elements from the *query case* as outputs. This approach provides the following three benefits:

- **Less noise and cost-effectiveness:** It reduces noise compared to using an existing retriever's results as gold query-target pairs (e.g., BM25) (Li et al., 2023a).

- **Balanced training on long-tail crimes:** It enables better sampling of rare case types than citation-based approaches.
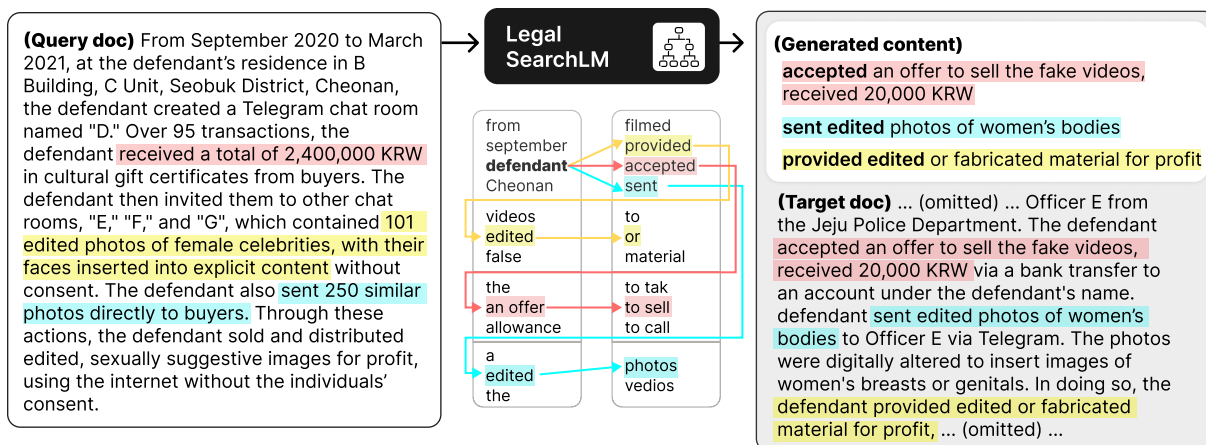
Figure 4: Inference process of LEGALSEARCHLM. Given a **Query doc** as input, LEGALSEARCHLM generates key legal elements expected to appear in the **Target doc** via *core-first-token-aware* constrained decoding over a prefix-indexed corpus (**Generated content**). Since the generated content is grounded in the corpus, it can be linked back to its source document, enabling retrieval.

- **Better generalization:** The model learns to reason over legal elements from a query case, aligning with inference-time conditions without relying on memorization.

3.2 Inference

During inference, LEGALSEARCHLM performs constrained beam decoding over the document index, ensuring that the generated content always exists in a document within the corpus (See Figure 4). For constrained decoding, we employ an FM-index based on the Burrows-Wheeler Transform (BWT) (Ferragina and Manzini, 2000). It enables efficient exact pattern matching via backward search in time linear to the pattern length. Detailed explanations on the FM-index are in Appendix A.1. The specific generation process is as follows.

Let $x_1, x_2, \ldots, x_n$ denote a generated token sequence obtained via constrained beam search. At each decoding step t, we maintain a beam of partial sequences $\{x_{<t}^{(1)}, x_{<t}^{(2)}, \ldots, x_{<t}^{(B)}\}$, where each hypothesis is extended based on the previously generated tokens $x_{<t}^{(i)}$. The candidate set $\mathcal{C}(x_{<t}^{(i)})$ represents the valid next tokens for the i-th hypothesis, constrained by FM-index. This ensures that all generated sequences present within the corpus.

$$
x_1^{(i)} = \arg\max_{x \in V} P(x \mid [\text{BOS}])
$$
$$
\text{for } i = 1, \ldots, B \tag{1}
$$
$$
x_t^{(i)} \in \text{Top-}k \left\{ P(x \mid x_{<t}^{(i)}) \,\Big|\, x \in \mathcal{C}(x_{<t}^{(i)}) \right\}
$$
$$
\text{for } t \geq 2 \tag{2}
$$

In our retrieval process, LEGALSEARCHLM captures key legal elements from the query case that are expected to appear in the target case, and begins generation with informative initial tokens to guide decoding toward relevant parts of the corpus.

4 Experimental Setup

4.1 Baselines

We evaluate a range of baselines, including traditional lexical matching and embedding-based methods trained on general or legal domains.

Lexical matching. (1) BM25, a strong baseline in the legal domain (Rosa et al., 2021), widely adopted by production-level legal RAG systems (Magesh et al., 2024).

General dual-encoders. (1) CONTRIEVER (Izacard et al., 2022), a dual-encoder model designed as a general-purpose retriever. We further adapt this model by training it on our legal corpus. (2) ME5 (Wang et al., 2024b), an open-source multilingual E5 (Wang et al., 2024a) embedding model that has achieved state-of-the-art retrieval performance. (3) OPENAI-EMBEDDING [1], a widely used commercial embedding model from OpenAI (`text-embedding-3-small`).

Legal dual-encoders. (1) SAILER (Li et al., 2023a), which achieves strong performance in the LCR task of the COLIEE 2023 competition (Goebel et al., 2024). Compared to CONTRIEVER,

---

[1] https://platform.openai.com/docs/guides/embeddings/embedding-models

Criminal Category	LEGAR BENCH$_{Standard}$ (P@5)							LEGAR BENCH$_{Stricter}$ (P@5)			
	LegalSearchLM	BM25	Contriever	SAILER	Hybrid	mE5	OpenAI-Embedding	LegalSearchLM	BM25	Contriever	SAILER
Fraud	0.74	0.53	0.57	0.64	0.66	0.63	0.65	0.33	0.33	0.03	0.24
Injury or Violence	0.62	0.50	0.42	0.66	0.68	0.57	0.56	0.35	0.35	0.01	0.23
Sexual crime	0.62	0.49	0.40	0.65	0.66	0.51	0.53	0.37	0.37	0.02	0.32
Finance or Insurance	0.72	0.56	0.60	0.64	0.68	0.68	0.76	0.32	0.32	0	0.20
Defamation or Insult	0.83	0.58	0.48	0.78	0.78	0.68	0.70	0.33	0.33	0	0.22
Drug	0.80	0.52	0.76	0.84	0.84	0.88	0.84	0.34	0.33	0	0.10
Murder	0.50	0.50	0.30	0.50	0.50	0.30	0.30	0.39	0.39	0	0.35
Traffic offenses	0.47	0.33	0.33	0.94	0.94	0.82	0.89	0.34	0.34	0.02	0.12
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮				
Total	**0.68**	0.51	0.48	0.62	0.65	0.57	0.58	**0.35**	0.34	0.01	0.22

Table 3: Results on LEGAR BENCH$_{Standard}$ and LEGAR BENCH$_{Stricter}$. We present results for the 8 crime categories shared by both evaluation sets, while the scores for the remaining 25 crime categories in LEGAR BENCH$_{Standard}$ are listed in Table B.1.
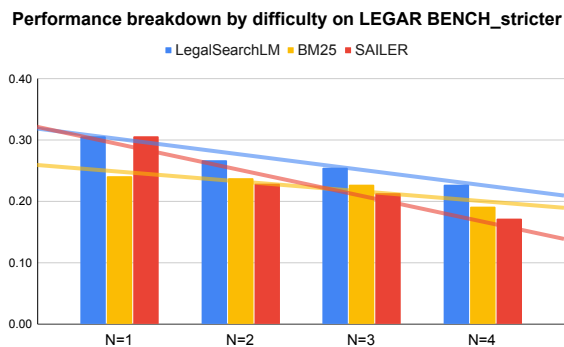

Figure 5: Performance on LEGAR BENCH$_{Stricter}$ by four different difficulties, where N represents the number of factors that should be matched. LEGALSEARCHLM achieves the best performance across all difficulty levels, demonstrating robustness in complex retrieval settings.

SAILER is pretrained using legal documents with section-level training loss (e.g., fact, interpretation (reasoning), and decision), enabling better legal document understanding. (2) HYBRID, obtained by averaging BM25 and SAILER scores.

5 Results and Analysis

Performance on LEGAR BENCH$_{Standard}$. An evaluation on the standard version, consisting of 411 various query types across 33 crime categories, demonstrates that LEGALSEARCHLM outperforms Contriever by 20%, BM25 by 17%, mE5 by 11%, OpenAI-Embedding by 10%, SAILER by 6%, and Hybrid by 3% (Table 3, bottom row). Specifically, it outperformed BM25 in 28 crime categories, Contriever across all categories, and SAILER in 21 cat-

egories. In Table 3, we provide the results for 8 out of 33 criminal categories for brevity. Full results are presented in Appendix B.1, where we also provide a comparison with the reranked model, KELLER.

Performance on LEGAR BENCH$_{Stricter}$. An evaluation on the stricter version, which includes 15,777 diverse query types across 8 crime categories, further demonstrates LEGALSEARCHLM's effectiveness in handling complex legal knowledge, achieving the highest performance in Table 3. BM25 excels at capturing fine-grained details through exact lexical matching, leading to stronger performance in LEGAR BENCH$_{Stricter}$ compared to embedding-based similarity search. LEGALSEARCHLM effectively captures both fine-grained details and legal semantic understanding, combining the strengths of both approaches. Further analysis is provided below.

Advantages of LEGALSEARCHLM over existing retrieval methods. In the LEGAL BENCH$_{Stricter}$ setting, we analyze retrieval performance across varying levels of retrieval difficulty, measured by the number of sub-factors N.

As shown in Figure 5, The performance of SAILER, an embedding-based retriever, relatively sharply degrades in difficult problems compared to LegalSearchLM, indicating that information is lost when vectorized. On the other hand, BM25, a lexical matching method, demonstrates no significant change by difficulty since it retrieves in a way that captures overlapping keywords without regard to legal element understanding. In contrast,
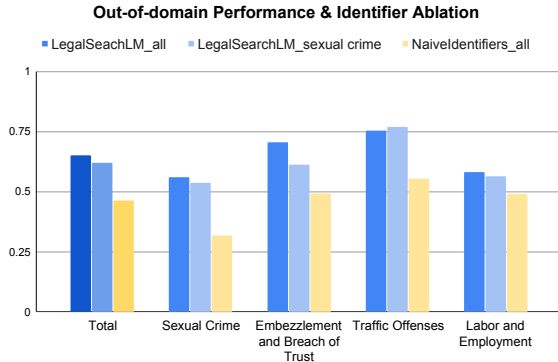
Out-of-domain Performance & Identifier Ablation

Figure 6: Performance on out-of-distribution. LegalSearchLM$_{all}$ is trained on all test categories; LegalSearchLM$_{sexual\ crime}$ on sexual crimes only. NaiveIdentifiers$_{all}$ is trained with random spans within the query case as supervision. Performance of LegalSearchLM$_{sexual\ crime}$ highlights its superior generalization ability.

LEGALSEARCHLM achieves the best performance in all difficulties by combining the strength of lexical matching in capturing fine-grained details with that of embedding-based retrieval in understanding semantics.

Generalization ability of LEGALSEARCHLM to unseen crime types. As legal professionals handle diverse cases, the generalizability to unseen criminal types is crucial in LCR. To evaluate this, we train LEGALSEARCHLM only using sexual crime data and test it on unseen domains (embezzlement and breach of trust, traffic offenses, and labor and employment). We compare the results with a generative retrieval model trained on all crimes but with naive identifiers. Figure 6 shows that LEGALSEARCHLM trained on sexual crime outperforms the model trained with naive identifiers by 15.66%, despite the latter is trained using in-domain data. Furthermore, the performance is almost on par with LEGALSEARCHLM trained on the full data. This demonstrates that the ability to effectively capture key legal elements is more beneficial than training on various datasets without carefully designed identifiers.

6   Related Works

6.1   Legal Case Retrieval Datasets

Legal Case Retrieval (LCR) is the task of retrieving target cases *relevant* to a query case (Feng et al., 2024; Ma et al., 2021). While some works define

relevant documents as one that is cited by the query (Shao et al., 2020; Li et al., 2023c), we focus on case similarity (Ma et al., 2021).

In LCR, obtaining large-scale data is challenging as the annotation requires legal expertise. Therefore, existing works often restrict the crime types or the retrieval pool's size (Li et al., 2023b; Ma et al., 2021). In contrast, this work successfully automated the annotation process while maintaining expert-level relevance judgments, eliminating the need for compromising data size and diversity.

In contrast, LEGAR BENCH$_{Standard}$ has effectively scaled the number of distinct crimes and the number of documents in the retrieval pool by using statutory provisions. Furthermore, LEGAR BENCH$_{Stricter}$ can evaluate the relevance based on expert-annotated legal factors on a large scale, which was not possible before.

6.2   Legal Case Retrievers

Earlier works on LCR have directly applied task-agnostic neural retrieval methods like cross-encoders (Xiao et al., 2021). However, recent works emphasize the specificity of the legal domain. For instance, SAILER (Li et al., 2023a) incorporates the document structure of legal cases during the pretraining, improving the embedding quality. KELLER and Elem4LCR first segment the case into atomic legal elements, and apply element-wise embedding similarity (KELLER) or cross-encoder scoring (Elem4LCR) between cases to obtain a fine-grained similarity score. LegalSearchLM proposes novel LCR-specific adjustments specialized for generative language models, while previous works resort to encoder-only models.

6.3   Generative Retrieval

Generative Retrieval (GR) initially emerged from GENRE (Cao et al., 2021), in which an encoder-decoder model retrieves a document by generating the title of the document from a given query.

Recent works explore identifiers based on specific document IDs (Tay et al., 2022; Mehta et al., 2023; Wang et al., 2023; Chen et al., 2024b; Zeng et al., 2023, 2024) or the document's *content*, which contain richer, finer-grained information. For instance, SEAL uses spans from the body text (Bevilacqua et al., 2022), which was also adopted by MINDER (Li et al., 2023f) and LTRGR (Li et al., 2023e) that combine spans, titles, and pseudo-queries. LEGALSEARCHLM is inspired

by content identifiers and incorporates domain-specific adaptations to better address legal retrieval tasks.

7 Conclusion

Legal case retrieval is a crucial task for legal practitioners, but resource scarcity and complex relevance judgment have hindered its application. To address these issues, we first construct LEGAR BENCH, a novel large-scale LCR dataset based on expert-defined relevance. LEGAR BENCH comes with two subsets, one with broader crime coverage and one with finer-grained relevance labels, supporting diverse applications. Next, we present LEGALSEARCHLM, the first generative retrieval model, that captures the core legal elements relevant to the given query case. LEGALSEARCHLM shows promising results in both versions of LEGAR BENCH, proving the potential of generative retrieval in LCR.

8 Limitations

In this dataset, we construct the largest benchmark in the legal case retrieval task, LEGAR BENCH. However, this dataset is restricted to the cases and statutes from the Korean legal system, which might limit its applicability beyond other jurisdictions and to non-Korean speakers. Furthermore, although legal experts were actively involved in defining the relevance criteria in LEGAR BENCH, they were not involved in the data point-wise verification of case-to-case relevance. As a result, there may be undetected noise in the dataset.

Acknowledgments

References

Michele Bevilacqua, Giuseppe Ottaviano, Patrick Lewis, Wen tau Yih, Sebastian Riedel, and Fabio Petroni. 2022. Autoregressive search engines: Generating substrings as document identifiers. *Preprint*, arXiv:2204.10628.

Fengyu Cai, Xinran Zhao, Tong Chen, Sihao Chen, Hongming Zhang, Iryna Gurevych, and Heinz Koeppl. 2024. MixGR: Enhancing retriever generalization for scientific domain through complementary granularity. *Preprint*, arXiv:2407.10691.

Nicola De Cao, Gautier Izacard, Sebastian Riedel, and Fabio Petroni. 2021. Autoregressive entity retrieval. *Preprint*, arXiv:2010.00904.

Tong Chen, Hongwei Wang, Sihao Chen, Wenhao Yu, Kaixin Ma, Xinran Zhao, Hongming Zhang, and Dong Yu. 2024a. Dense x retrieval: What retrieval granularity should we use? In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 15159–15177, Miami, Florida, USA. Association for Computational Linguistics.

Tong Chen, Hongwei Wang, Sihao Chen, Wenhao Yu, Kaixin Ma, Xinran Zhao, Hongming Zhang, and Dong Yu. 2024b. Dense x retrieval: What retrieval granularity should we use? *Preprint*, arXiv:2312.06648.

Chenlong Deng, Zhicheng Dou, Yujia Zhou, Peitian Zhang, and Kelong Mao. 2024a. An element is worth a thousand words: Enhancing legal case retrieval by incorporating legal elements. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 2354–2365, Bangkok, Thailand. Association for Computational Linguistics.

Chenlong Deng, Kelong Mao, and Zhicheng Dou. 2024b. Learning interpretable legal case retrieval via knowledge-guided case reformulation. *Preprint*, arXiv:2406.19760.

Yi Feng, Chuanyi Li, and Vincent Ng. 2024. Legal case retrieval: A survey of the state of the art. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6472–6485, Bangkok, Thailand. Association for Computational Linguistics.

P. Ferragina and G. Manzini. 2000. Opportunistic data structures with applications. In *Proceedings 41st Annual Symposium on Foundations of Computer Science*, pages 390–398.

Randy Goebel, Yoshinobu Kano, Mi-Young Kim, Juliano Rabelo, Ken Satoh, and Masaharu Yoshioka. 2024. Overview and discussion of the competition on legal information, extraction/entailment (coliee) 2023. *The Review of Socionetwork Strategies*, 18(1):27–47.

Wonseok Hwang, Saehee Eom, Hanuhl Lee, Hai Jin Park, and Minjoon Seo. 2022. Data-efficient end-to-end information extraction for statistical legal analysis. In *Proceedings of the Natural Legal Language Processing Workshop 2022*, pages 143–152, Abu Dhabi, United Arab Emirates (Hybrid). Association for Computational Linguistics.

Gautier Izacard, Mathilde Caron, Lucas Hosseini, Sebastian Riedel, Piotr Bojanowski, Armand Joulin, and Edouard Grave. 2022. Unsupervised dense information retrieval with contrastive learning. *Preprint*, arXiv:2112.09118.

Chaeeun Kim, Soyoung Yoon, Hyunji Lee, Joel Jang, Sohee Yang, and Minjoon Seo. 2024. Exploring the practicality of generative retrieval on dynamic corpora. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 13616–13633, Miami, Florida, USA. Association for Computational Linguistics.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising sequence-to-sequence pretraining for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7871–7880, Online. Association for Computational Linguistics.

Haitao Li, Qingyao Ai, Jia Chen, Qian Dong, Yueyue Wu, Yiqun Liu, Chong Chen, and Qi Tian. 2023a. Sailer: Structure-aware pre-trained language model for legal case retrieval. *Preprint*, arXiv:2304.11370.

Haitao Li, Yunqiu Shao, Yueyue Wu, Qingyao Ai, Yixiao Ma, and Yiqun Liu. 2023b. Lecardv2: A large-scale chinese legal case retrieval dataset. *Preprint*, arXiv:2310.17609.

Haitao Li, Weihang Su, Changyue Wang, Yueyue Wu, Qingyao Ai, and Yiqun Liu. 2023c. Thuir@coliee 2023: Incorporating structural knowledge into pretrained language models for legal case retrieval. *Preprint*, arXiv:2305.06812.

Qingquan Li, Yiran Hu, Feng Yao, Chaojun Xiao, Zhiyuan Liu, Maosong Sun, and Weixing Shen. 2023d. Muser: A multi-view similar case retrieval dataset. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, CIKM '23, page 5336–5340. ACM.

Xiaoxi Li, Jiajie Jin, Yujia Zhou, Yuyao Zhang, Peitian Zhang, Yutao Zhu, and Zhicheng Dou. 2024. From matching to generation: A survey on generative information retrieval. *Preprint*, arXiv:2404.14851.

Yongqi Li, Nan Yang, Liang Wang, Furu Wei, and Wenjie Li. 2023e. Learning to rank in generative retrieval. *Preprint*, arXiv:2306.15222.

Yongqi Li, Nan Yang, Liang Wang, Furu Wei, and Wenjie Li. 2023f. Multiview identifiers enhanced generative retrieval. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6636–6648, Toronto, Canada. Association for Computational Linguistics.

Yixiao Ma, Yunqiu Shao, Yueyue Wu, Yiqun Liu, Ruizhe Zhang, Min Zhang, and Shaoping Ma. 2021. Lecard: A legal case retrieval dataset for chinese law system. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '21, page 2342–2348, New York, NY, USA. Association for Computing Machinery.

Varun Magesh, Faiz Surani, Matthew Dahl, Mirac Suzgun, Christopher D. Manning, and Daniel E. Ho. 2024. Hallucination-free? assessing the reliability of leading ai legal research tools. *Preprint*, arXiv:2405.20362.

Sanket Vaibhav Mehta, Jai Gupta, Yi Tay, Mostafa Dehghani, Vinh Q. Tran, Jinfeng Rao, Marc Najork, Emma Strubell, and Donald Metzler. 2023. Dsi++: Updating transformer memory with new documents. *Preprint*, arXiv:2212.09744.

Sewon Min, Kalpesh Krishna, Xinxi Lyu, Mike Lewis, Wen tau Yih, Pang Wei Koh, Mohit Iyyer, Luke Zettlemoyer, and Hannaneh Hajishirzi. 2023. Factscore: Fine-grained atomic evaluation of factual precision in long form text generation. *Preprint*, arXiv:2305.14251.

Guilherme Moraes Rosa, Ruan Chaves Rodrigues, Roberto Lotufo, and Rodrigo Nogueira. 2021. Yes, bm25 is a strong baseline for legal case retrieval. *Preprint*, arXiv:2105.05686.

Yunqiu Shao, Bulou Liu, Jiaxin Mao, Yiqun Liu, Min Zhang, and Shaoping Ma. 2020. Thuir@coliee-2020: Leveraging semantic understanding and exact matching for legal case retrieval and entailment. *Preprint*, arXiv:2012.13102.

Weihang Su, Qingyao Ai, Yueyue Wu, Yixiao Ma, Haitao Li, Yiqun Liu, Zhijing Wu, and Min Zhang. 2024. Caseformer: Pre-training for legal case retrieval based on inter-case distinctions. *Preprint*, arXiv:2311.00333.

Yi Tay, Vinh Q. Tran, Mostafa Dehghani, Jianmo Ni, Dara Bahri, Harsh Mehta, Zhen Qin, Kai Hui, Zhe Zhao, Jai Gupta, Tal Schuster, William W. Cohen, and Donald Metzler. 2022. Transformer memory as a differentiable search index. *Preprint*, arXiv:2202.06991.

Liang Wang, Nan Yang, Xiaolong Huang, Binxing Jiao, Linjun Yang, Daxin Jiang, Rangan Majumder, and Furu Wei. 2024a. Text embeddings by weakly-supervised contrastive pre-training. *Preprint*, arXiv:2212.03533.

Liang Wang, Nan Yang, Xiaolong Huang, Linjun Yang, Rangan Majumder, and Furu Wei. 2024b. Multilingual e5 text embeddings: A technical report. *Preprint*, arXiv:2402.05672.

Yujing Wang, Yingyan Hou, Haonan Wang, Ziming Miao, Shibin Wu, Hao Sun, Qi Chen, Yuqing Xia, Chengmin Chi, Guoshuai Zhao, Zheng Liu, Xing Xie, Hao Allen Sun, Weiwei Deng, Qi Zhang, and Mao Yang. 2023. A neural corpus indexer for document retrieval. *Preprint*, arXiv:2206.02743.

Chaojun Xiao, Xueyu Hu, Zhiyuan Liu, Cunchao Tu, and Maosong Sun. 2021. Lawformer: A pre-trained language model for chinese legal long documents. *Preprint*, arXiv:2105.03887.

Chaojun Xiao, Haoxi Zhong, Zhipeng Guo, Cunchao Tu, Zhiyuan Liu, Maosong Sun, Yansong Feng, Xianpei Han, Zhen Hu, Heng Wang, and Jianfeng Xu. 2018. Cail2018: A large-scale legal dataset for judgment prediction. *Preprint*, arXiv:1807.02478.

Linting Xue, Noah Constant, Adam Roberts, Mihir Kale, Rami Al-Rfou, Aditya Siddhant, Aditya Barua, and Colin Raffel. 2021. mT5: A massively multilingual pre-trained text-to-text transformer. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 483–498, Online. Association for Computational Linguistics.

Hansi Zeng, Chen Luo, Bowen Jin, Sheikh Muhammad Sarwar, Tianxin Wei, and Hamed Zamani. 2023. Scalable and effective generative information retrieval. *Preprint*, arXiv:2311.09134.

Hansi Zeng, Chen Luo, and Hamed Zamani. 2024. Planning ahead in generative retrieval: Guiding autoregressive generation through simultaneous decoding. *Preprint*, arXiv:2404.14600.

Kun Zhang, Chong Chen, Yuanzhuo Wang, Qi Tian, and Long Bai. 2023. Cfgl-lcr: A counterfactual graph learning framework for legal case retrieval. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '23, page 3332–3341, New York, NY, USA. Association for Computing Machinery.

A Inference Details

A.1 Background on FM-index

FM-index. The FM-index is a compressed full-text index built on top of the Burrows–Wheeler Transform (BWT). Given a text T terminated by a unique sentinel (commonly "$"), the BWT is obtained by generating all cyclic rotations of T, sorting them lexicographically, and extracting the last column of the resulting matrix.

Data structures. The FM-index augments the BWT with two auxiliary structures. The array C stores, for each character c, the number of characters in T that are lexicographically smaller than c. The table $\mathrm{Occ}(c, i)$ counts how many occurrences of c appear in the BWT up to position i. Together, these enable efficient navigation between the first and last columns of the BWT.

Backward search. Pattern matching proceeds right-to-left: we start with the entire text as a candidate range and, at each step, use C and Occ to restrict the range to suffixes consistent with the processed part of the pattern. If the range becomes empty, the pattern does not occur; otherwise, the final range corresponds to all matches. This procedure runs in time proportional to the pattern length.

Use in constrained decoding. In our setting, the FM-index acts as a prefix-tree–like constraint over the corpus, restricting decoding to continuations that appear exactly in T. As a result, early token choices are critical, since they determine whether a valid continuation path remains available.

B Result Details

B.1 Full results on LEGAR BENCH*Standard*

We provide the complete results on LEGAR BENCH*Standard* across all 33 criminal categories in Table 4. Keep in mind that KELLER, which focuses on reranking, leverages passage-level retrieval and make multiple inferences per case using majority voting (*MaxSum*). This setup differs from our model and other baselines, making direct comparisons difficult. We include KELLER as a reference of a reranked model. All results are from single-run experiments.

C Benchmark Details

C.1 Collaboration with legal experts

Annotation of LEGAR BENCH*Stricter* requires a significant amount of legal expert annotation. For instance, determining the critical factors that determine the applicability of a specific statute requires extensive knowledge of criminal law, and determining the range of inherently continuous values (*e.g.* severity of an injury) that are similarly treated in practice requires strong expertise in practicing criminal law.

For expert annotation, we hired five Korean lawyers specialized in the Criminal Act to construct our LEGAR BENCH*Standard* and LEGAR BENCH*Stricter*. The lawyers were instructed to organize high-coverage categories that encompass most criminal offenses and to label subcategories of cases based on charge titles. Additionally, for the stricter version, they listed relevant factors (factual details) for each specific charge in LEGAR BENCH*Standard*. Lawyers spent a total 70 hours for the annotation task, and the compensation was approximately \$250/hour during the whole process.

C.2 Statistics of LEGAR BENCH*Standard*

Table 5 presents a criminal typology that includes 33 major categories of criminal offenses. Each category is classified in detail based on charge titles and statutes. The number of standard groups for each category is listed under *#of Standard Group*, while the number of unique case documents mapped to each group is listed under *#of Cases*. The total number of standard groups is 411, including 1,052,506 unique cases, which constitute 85.79% of the entire corpus (1,226,814 cases). This figure underscores the broad coverage of our benchmark across a wide range of types of criminal offenses.

C.3 Statistics of LEGAR BENCH*Stricter*

Table 6 shows statistics of stricter groups in 8 criminal categories. The number of stricter groups for each category is listed under *#of Stricter Group*, while the number of unique case documents mapped to each group is listed under *#of Cases*.

C.4 Total crime types in 33 categories for query cases.

Traffic offenses. In Table 25.

Sexual crime. In Table 8

| | LEGAR BENCH$_{Standard}$ (Precision@5) | | | | | | | |
Criminal Category	LegalSeachLM(Ours)	BM25	Contriever	SAILER	Hybrid	mE5	OpenAI-Embedding	KELLER*
[Total]	**[0.68]**	[0.51]	[0.48]	[0.62]	[0.65]	[0.57]	[0.58]	[0.70]
Traffic offenses	0.75	0.60	0.72	0.94	0.94	0.82	0.89	0.89
Fraud	0.74	0.53	0.57	0.64	0.66	0.63	0.65	0.90
Injury or Violence	0.62	0.50	0.42	0.66	0.68	0.57	0.56	0.70
Sexual crime	0.52	0.49	0.40	0.65	0.66	0.51	0.53	0.73
Theft or Robbery	0.65	0.48	0.41	0.60	0.63	0.48	0.47	0.69
Obstruction of Business	0.75	0.58	0.40	0.74	0.74	0.65	0.57	0.94
Embezzlement or Breach of trust	0.67	0.64	0.49	0.56	0.60	0.55	0.60	0.81
Destruction	0.64	0.64	0.48	0.72	0.68	0.48	0.60	0.92
Finance or Insurance	0.72	0.56	0.60	0.64	0.68	0.68	0.76	0.92
Threat	0.60	0.60	0.47	0.60	0.64	0.62	0.75	0.87
Defamation or Insult	0.83	0.58	0.48	0.78	0.78	0.68	0.70	0.80
Drug	0.80	0.52	0.76	0.84	0.84	0.88	0.84	0.80
Criminal trespass	0.73	0.63	0.51	0.80	0.80	0.65	0.63	0.89
Gambling	0.74	0.54	0.63	0.71	0.74	0.89	0.74	0.97
Negligent homicide and injury	0.37	0.27	0.30	0.57	0.57	0.43	0.40	0.50
Obstruction of right	0.80	0.52	0.52	0.52	0.56	0.40	0.52	0.92
Child abuse or School violence	0.64	0.48	0.38	0.50	0.52	0.62	0.50	0.54
Medical or Food drug	0.40	0.35	0.22	0.11	0.24	0.25	0.22	0.14
Murder	0.50	0.50	0.30	0.50	0.50	0.30	0.30	0.60
Corporation	0.47	0.33	0.33	0.33	0.47	0.33	0.27	0.40
Bribery	0.47	0.60	0.40	0.80	0.80	0.47	0.53	0.80
Car	0.80	0.60	0.60	0.60	0.60	0.60	0.70	1.00
Labor or Employment	0.53	0.51	0.40	0.58	0.64	0.56	0.53	0.55
Industrial or Serious accidents	0.45	0.45	0.20	0.40	0.50	0.50	0.30	0.25
Military duty or law	0.70	0.50	0.50	0.60	0.60	0.60	0.60	0.50
Consumer or Fair trade	1.00	0.60	0.80	0.40	0.40	0.60	0.40	0.40
Arrest or Detention	1.00	0.80	0.40	0.80	0.80	0.40	0.60	1.00
Intellectual property	0.60	0.33	0.67	0.80	0.80	0.80	0.73	0.73
IT or Privacy	1.00	0.60	0.80	1.00	1.00	0.80	0.60	1.00
Misdemeanor	0.40	0.20	0.20	0.40	0.40	0.40	0.20	0.40
Sexual norms	0.80	0.20	0.20	0.20	0.20	0.40	1.00	0
Tax, Administ, Const law	0.81	0.61	0.71	0.91	0.91	0.74	0.71	0.76
Other criminal offenses	0.72	0.58	0.42	0.70	0.74	0.66	0.70	0.88

Table 4: Full results of LegalSearchLM and the baselines on LEGAR BENCH$_{Standard}$ across all 33 criminal categories.

Fraud. In Table 9.

Injury and Violence. In Table 14.

Theft and Robbery. In Table 10.

Embezzlement. In Table 11.

Destruction. In Table 12.

Finance and Insurance. In Table 32.

Threat. In Table 17.

Crimes against Reputation. In Table 18.

Drug. In Table 22.

Gambling. In Table 23.

Negligent homicide and injury. In Table 15.

Obstruction of rights. In Table 21.

Crime categories	# of Standard group	# of Cases
Traffic offenses	13	319,527
Fraud	21	181,703
Injury or Violence	31	146,764
Sexual crime	132	104,919
Theft or Robbery	38	74,772
Obstruction of Business	13	74,722
Embezzlement or Breach of trust	15	39,835
Destruction	5	39,595
Finance or Insurance	5	32,944
Threat	11	27,496
Defamation or Insult	8	27,278
Drug	5	26,066
Criminal trespass	15	24,856
Gambling	7	11,091
Negligent homicide and injury	6	7,384
Obstruction of right	5	6,749
Child abuse or School violence	10	5,756
Medical or Food drug	11	98
Murder	2	4,306
Corporation	3	1,195
Bribery	3	1,638
Car	2	20,882
Labor or Employment	11	12,647
Industrial or Serious accidents	4	198
Military duty or law	2	9,300
Consumer or Fair trade	1	128
Arrest or Detention	1	6
Intellectual property	3	3,927
IT or Privacy	2	2,311
Misdemeanor	1	6,476
Sexual norms	1	4,140
Tax, Administ, Const law	14	40,890
Other criminal offenses	10	23,211
Total	**411**	**1,052,506**

Table 5: Statistics of Crime typology and Standard version of LEGAR BENCH. The total number of cases is reported as a unique count, excluding duplicates from cases classified under multiple categories $1,347,962 \rightarrow 1,052,506$.

Crime categories	# of Stricter group	# of Cases
Fraud	8	325
Injury or Violence	19	308
Sexual crime	111	1,061
Finance or Insurance	1	28
Defamation or Insult	6	253
Drug	4	37
Murder	2	8
Traffic offenses	9	330
Total	**160**	**2,350**

Table 6: Statistics of Stricter version of LEGAR BENCH.

Crimes against children and School violence. In Table 26.

Medical and Food drug. In Table 30.

Murder. In Table 13.

Corporation. In Table 27.

Bribery. In Table 37.

Labor and Employment. In Table 28.

Fair trade. In Table 31.

Arrest and Detention. In Table 16.

Other criminal offenses. In Table 24.

Car-related offenses. In Table 40.

Home invasion. In Table 20.

Industrial accident or Serious accident. In Table 29.

Intellectual property rights. In Table 38.

IT or Privacy. In Table 36.

Military duty. In Table 39.

Misdemeanor. In Table 35.

Obstruction of business. In Table 19.

Sexual morality. In Table 34.

Tax or Administrative or Constitutional Law. In Table 33.

C.5 Full list of the stricter relevance group

LEGAR BENCH$_{Stricter}$ further divides LEGAR BENCH$_{Standard}$ categories based on different factual details of a criminal case that do not affect the type of charge, but might affect the final judgment (guilty or innocent) or the sentence *e.g.* information about defendant/victims, methods, consequences, and claims made in court. Also, it provides a comprehensive list of possible options for each factor. The options are primarily based on the official sentencing guidelines from the Sentencing Commission of the Supreme Court of Korea, and annual crime statistics reports published by government/academic authorities including the Supreme Prosecutor's Office and the Korean Institute of Criminology. However, these lists are often insufficient to express existing cases, especially the defendant's claims (*e.g.*, a defendant convicted of assault might claim that the act was due to self-defense, pleading for innocence). Identifying such factors heavily relies on deep understanding and expertise in practicing law. Hence, the lawyers were instructed to

add factors and options that are frequent and important in practice but not mentioned in the official documents.

Previous work in identifying such factors in the Korean Criminal Act Hwang et al. (2022) includes only 11 unique factors across 4 crime categories focusing only on facts, while this work adds 102 unique factors (including 39 defendant claims) across 8 categories.

C.6 Prompt template for LEGAR BENCH*Stricter* annotation

The example annotation templates are shown on the following page. The blue box presents the original Korean version, and the pink box presents the English version provided for reference.

D Implementation Details

All models are trained using 8 * A100 80GB GPUs.

LegalSearchLM. To develop our SearchLM based on an autoregressive language model, we take the mt5-base pretrained model (Xue et al., 2021) and train it on 170K cases for a single epoch. We create a training dataset with a maximum of 15 query case-element pairs and 5 element-element pairs.

Contriever. We select Contriever as a representative model for retrieval in the general domain. We perform unsupervised training on the bert-base-multilingual-cased pretrained model with 170K cases for 10 epochs. Following the results in their work, we use the MoCo method during training rather than in-batch.

SAILER. We implement SAILER as a representative model for retrieval in the legal domain. Following their paper, we pretrain the bert-base-multilingual-cased model on facts, interpretations, and decisions of 1.2M cases for a single epoch, using the same configuration as in SAILER. The pretrained model is then fine-tuned for a single epoch with positive and negative samples, adjusting the learning rate from the default 5e-6 to 5e-5. We retrieve 100 related cases using BM25 over the 170K cases, selecting those with the same case name as positive samples and others as negative. To ensure comparability with other baselines, we use 5 positive and 5 negative cases per query.

KELLER. We implement KELLER based on the code from the official repository[3]. To prepare the retrieval pool, we first process the same 1.2M cases used in implementing SAILER. To further separate cases into subfactual levels, we use GPT-4o. Each subfact is labeled with its criminal type using either their subheadings or regular expressions. This results in 1.1M cases.

For training, we use the same query cases from SearchLM. Following the same process, we separate and label each subfact, resulting in 143K cases. We prepare ground truth document cases by matching cases that include the same subfacts criminal type as the query case. Among the matching cases, we use the top 10 document cases based on BM25 scores. This results in total of 820k (query case, document case) pairs. The model is trained on the resulting dataset for 1 epoch under the same condition with the original Keller paper (batch size: 128, learning rate: 1e-5, optimizer: AdamW).

E Licenses and intended use

Korean legal cases are not protected by the Korean Copyright Act. mT5 (Xue et al., 2021), the base model of LEGALSEARCHLM, is disclosed with Apache 2.0 license that permits free academic use.

Korean legal cases are fully anonymized when disclosed by the Korean court. However, we do not censor potentially offensive content, including descriptions about violent and sexual crimes, as they constitute the core content of legal cases.

---

[3]https://github.com/ChenlongDeng/KELLER

Example prompt for a specific d1–d2 case (Korean)

Input으로 주어지는 판결문은 d1 카테고리의 d2 범죄에 대한 내용입니다. 아래 각 항목의 분류 기준을 기반으로 항목별 분류 번호를 제공하세요. 모든 항목에 대해 누락없이 분류 번호를 제공하세요. 단, 답변은 Output 예시처럼 항목과 분류 번호로 구성된 튜플이 나열된 리스트 형식으로 작성하세요.

"d1": "명예훼손및모욕",
"d2": "허위적시명예훼손",
"description":
"<형사공통>
- 항목: 자수여부
- 항목 설명: 피고인이 범행 이후 자수하였는지 여부
- 분류 기준:
1. 피고인이 자수함 2. 피고인이 자수하지 않음

- 항목: 심신미약여부
- 항목 설명: 피고인이 범행 당시 심신미약이었는지 여부
- 분류 기준:
1. 해당 쟁점이 다루어지고, 인정됨 (사실관계에 '심신미약 상태로...'의 표현이 있는 경우 포함) 2. 해당 쟁점이 다루어지고, 인정되지 않음 3. 해당 쟁점이 다뤄지지 않음

- 항목: 심신상실여부
- 항목 설명: 피고인이 범행 당시 심신상실이었는지 여부
- 분류 기준:
1. 해당 쟁점이 다루어지고, 인정됨 (사실관계에 '심신상실 상태로...'의 표현이 있는 경우 포함) 2. 해당 쟁점이 다루어지고, 인정되지 않음 3. 해당 쟁점이 다뤄지지 않음

<피해자범죄>
- 항목: 피해자수
- 항목 설명: 피해자의 수
- 분류 기준:
1. 1명 2. 2명 3. 3명 이상

- 항목: 피해자와의관계
- 항목 설명: 피고인과 피해자와의 관계
- 분류 기준:
1. 연인 2. 부부 3. 피해자가 피고인의 부모 4. 피해자가 피고인의 자녀 5. 5촌 이내 친족 6. 친구/지인 7. 피해자가 피고인의 피감독자/피보호자 (교사/제자, 고용인/피고용인 등) 8. 동료직원 9. 클럽, 랜덤채팅 등 일회성 만남 10. 불특정 다수 대상 범행 (피해자와 가해자가 모르는 사이인 경우)

<명예훼손>
- 항목: 명예훼손유형
- 항목 설명: 명예훼손의 문제가 된 발언의 유형
- 분류 기준:
1. 과거 전과 기록 2. 직업적 명예와 관한 사실 (비리, 중대한 업무상 과실, 업무능력에 대한 평가 등) 3. 범죄행위 등과 관련된 의혹 제기 4. 치정/가족사/사생활 등 5. 기타

- 항목: 명예훼손매체유형
- 항목 설명: 명예훼손의 매체
- 분류 기준:
1. 일상적인 개인 간 대화 2. 연설/토론회 등 공적 대화 3. 벽보, 포스터, 공고문 등 공개된 문서 4. 정보통신망-인터넷 뉴스 등 공적 매체 5. 정보통신망-SNS, 커뮤니티, 채팅, 게임, 유튜브 등 개인 간의 대화 6. 출판물-신문, 잡지 등 7. 출판물-서적, 논문 등 8. 출판물-TV/라디오 등 매체

- 항목: 피해자유형
- 항목 설명: 피해자가 자연인인지 법인인지 여부
- 분류 기준:
1. 자연인 2. 법인 3. 법인격없는단체 (공중 등)

- 항목: 구체적사실의적시
- 항목 설명: 명예훼손죄가 성립하기 위한 구체적인 사실이 적시되었는지 여부가 쟁점인 경우
- 분류 기준:
1. 해당 쟁점이 다루어지고, 인정됨 2. 해당 쟁점이 다루어지고, 인정되지 않음 3. 해당 쟁점이 다뤄지지 않음

- 항목: 공연성전파가능성
- 항목 설명: 불특정 또는 다수인이 인식할 수 있는 상태 (공연성) 를 만족하는지, 제3자에게 전파될 가능성이 있는지가 쟁점인 경우
- 분류 기준:
1. 해당 쟁점이 다루어지고, 인정됨 2. 해당 쟁점이 다루어지고, 인정되지 않음 3. 해당 쟁점이 다뤄지지 않음

- 항목: 피해자특정성
- 항목 설명: 명예훼손만의 내용으로 피해자가 특정되는지 여부
- 분류 기준:
1. 해당 쟁점이 다루어지고, 인정됨 2. 해당 쟁점이 다루어지고, 인정되지 않음 3. 해당 쟁점이 다뤄지지 않음

- 항목: 사회적평가저해여부
- 항목 설명: 명예훼손의 내용이 피해자의 명예를 훼손하여 사회적 평가의 저해로 이어지는 사안인지가 쟁점이 된 경우
- 분류 기준:
1. 해당 쟁점이 다루어지고, 인정됨 2. 해당 쟁점이 다루어지고, 인정되지 않음 3. 해당 쟁점이 다뤄지지 않음

- 항목: 공익위법성조각정당행위
- 항목 설명: 명예훼손의 내용이 진실한 사실로서 오직 공공의 이익을 목적으로 하므로 위법성이 조각되거나, 언론사 등의 적법한 업무로서 정당한 행위인지 여부가 쟁점이 되는 경우
- 분류 기준:
1. 해당 쟁점이 다루어지고, 인정됨 2. 해당 쟁점이 다루어지고, 인정되지 않음 3. 해당 쟁점이 다뤄지지 않음"

Output 예시: [('피해자와의관계', 6), ('명예훼손유형', 2), ('명예훼손의 매체', 5), ('공연성전파가능성', 1), ('공익위법성조각정당행위', 3)]

Input: #d1카테고리 (명예훼손및모욕) d2범죄 (허위적시명예훼손)에 해당되는 각 판결문 예시 (생략)

Example prompt for a specific d1–d2 case (English)

The judgment provided as input concerns an offense in category d1 and specific crime type d2. Based on the classification criteria for each item below, provide the classification number for each item. Provide a classification number for every item without omission. The answer must be written in a list format consisting of tuples of (item, classification number), as in the Output example.

"d1": "Defamation and Insult",
"d2": "Defamation by False Facts",
"description":
<Criminal (Common)>
- Item: Voluntary surrender
- Item description: Whether the defendant voluntarily surrendered after the offense
- Classification criteria:
1. The defendant voluntarily surrendered 2. The defendant did not voluntarily surrender

- Item: Diminished capacity
- Item description: Whether the defendant was in a state of diminished capacity at the time of the offense
- Classification criteria:
1. The issue is addressed and recognized (including cases where the facts state 'in a state of diminished capacity...') 2. The issue is addressed but not recognized 3. The issue is not addressed

- Item: Insanity
- Item description: Whether the defendant was insane at the time of the offense
- Classification criteria:
1. The issue is addressed and recognized (including cases where the facts state 'in a state of insanity...') 2. The issue is addressed but not recognized 3. The issue is not addressed

<Victim-Related Offense>
- Item: Number of victims
- Item description: Number of victims
- Classification criteria:
1. 1 victim 2. 2 victims 3. 3 or more victims

- Item: Relationship with the victim
- Item description: Relationship between the defendant and the victim
- Classification criteria:
1. Romantic partner 2. Spouses 3. The victim is the defendant's parent 4. The victim is the defendant's child 5. Relative within the fifth degree of kinship 6. Friend/acquaintance 7. The victim is the defendant's supervisee/ward (teacher/student, employer/employee, etc.) 8. Coworker 9. One-off encounter (club, random chat, etc.) 10. Offense against unspecified persons (the perpetrator and victim are strangers)

<Defamation>
- Item: Type of defamatory statement
- Item description: The type of statement at issue for defamation
- Classification criteria:
1. Prior criminal record 2. Facts concerning professional reputation (corruption, serious professional negligence, evaluation of job ability, etc.) 3. Allegations related to criminal conduct, etc. 4. Romantic/family/private life, etc. 5. Other

- Item: Medium of defamation
- Item description: The medium of the defamatory act
- Classification criteria:
1. Everyday private conversation 2. Public speech/discussion (e.g., lecture, debate) 3. Publicly posted documents (posters, notices, etc.) 4. Information network – internet news and other public media 5. Information network – interpersonal channels such as SNS, communities, chat, games, YouTube, etc. 6. Publications – newspapers, magazines, etc. 7. Publications – books, academic papers, etc. 8. Publications – TV/radio, etc.

- Item: Type of victim
- Item description: Whether the victim is a natural person or a legal entity
- Classification criteria:
1. Natural person 2. Legal entity 3. Unincorporated association (the public, etc.)

- Item: Statement of specific facts
- Item description: When it is at issue whether specific facts were stated, which is required for defamation to be established
- Classification criteria:
1. The issue is addressed and recognized 2. The issue is addressed but not recognized 3. The issue is not addressed

- Item: Publicity / possibility of dissemination
- Item description: When it is at issue whether the requirement of publicity (recognizable by unspecified or many persons) is satisfied and whether there is a possibility of dissemination to third parties
- Classification criteria:
1. The issue is addressed and recognized 2. The issue is addressed but not recognized 3. The issue is not addressed

- Item: Specificity of the victim
- Item description: Whether the victim is identifiable solely from the content of the defamatory statement
- Classification criteria:
1. The issue is addressed and recognized 2. The issue is addressed but not recognized 3. The issue is not addressed

- Item: Impairment of social reputation
- Item description: When it is at issue whether the content of the defamation harms the victim's reputation and leads to impairment of social evaluation
- Classification criteria:
1. The issue is addressed and recognized 2. The issue is addressed but not recognized 3. The issue is not addressed

- Item: Public-interest / justifiable act (illegality exemption)
- Item description: When it is at issue whether the content of the defamation consists of true facts for the sole purpose of the public interest, thereby excluding illegality, or constitutes a justifiable act as part of legitimate activities such as those of the press
- Classification criteria:
1. The issue is addressed and recognized 2. The issue is addressed but not recognized 3. The issue is not addressed

Output example: [('Relationship with the victim', 6), ('Type of defamatory statement', 2), ('Medium of defamation', 5), ('Publicity / possibility of dissemination', 1), ('Public-interest / justifiable act (illegality exemption)', 3)]

Input: (omitted) example case corresponding to each d1-d2 pair

Crime categories	Factors(# Options)
Traffic offenses	Traffic accident type(6), Traffic accident time(2), Automobile type(3), Road type(4), Gross negligence type(18), Automobile accident insurance(3), Malpractice?(3), Hit-and-run type(3), Hit-and-run loss type(2), Aided victim?(3), Not aware of accident?(3), Blood alcohol level(3), Driving distance(4), Necessity?(3), Not driving?(3), Absorption phase?(3), Excessive extrapolation?(3), Driving without license type(5), Not aware of license suspension(3), Not aware of invalidation(3), Injury severity(8), Injury?(3), Number of victims(3), Defendant-victim relation(10), Surrender(2), Defendant feeble-minded?(3), Defendant insanity?(3), Reason not reaching consummation(4), Reached consummation?(3)
Fraud	Fraud type(14), No intent for pecuniary advantage?(3), No intent to defraud?(3), Profit(12), Defendant feeble-minded?(3), Defendant insanity?(3)
Injury or Violence	Two-way assault(2), Motivation(7), Intent to injure?(3), Self-defense?(3), Assault method(9), Injury severity(8), Injury?(3), Special crime type(2), Number of accomplices(5), Dangerous weapon?(3), Time between injury and death(4), Injury direct cause of death?(3), Surrender(2), Defendant feeble-minded?(3), Defendant insanity?(3)
Sexual crime	Sexual assault location(6), Victim age(4), Victim disability(2), Defendant under influence(3), Victim under influence(3), Consent?(3), Intercourse type(4), Incident act type(4), Incident act by blitz(2), Victim sexual shame(3), Inability to resist cause(5), Aware of inability to resist?(3), Aware of victim's age under 13?(3), Aware of victim's age under 16?(3), Fraudulence/influence type(7), Victim under influence?(3), Covert photography filming/distribution type(7), Number of covert photography(4), Profit(4), Obscene communication medium(4), Obscene communication content(6), Object of sexual satisfaction(2), Reached the victim?(3), Assault/threat type(6), Assault method(9), Injury severity(8), Injury?(3), Special crime type(2), Number of accomplices(5), Dangerous weapon?(3), Time between injury and death(4), Injury direct cause of death?(3), No intent to defraud?(3), Number of victims(3), Defendant-victim relation(10), Surrender(2), Defendant feeble-minded?(3), Defendant insanity?(3), Reason not reaching consummation(4), Reached consummation?(3)
Finance or Insurance [2]	Insurance fraud type(5), No intent for pecuniary advantage?(3), No intent to defraud?(3), Profit(12), Surrender(2), Defendant feeble-minded?(3), Defendant insanity?(3), Reason not reaching consummation(4), Reached consummation?(3)
Defamation or Insult	Defamation content(5), Defamation medium(8), Insult content(4), Victim type(3), Alleged facts?, Publicly alleged?(3), Can specify victim?(3), Defaming the social status?(3), Justified(3), Number of victims(3), Defendant-victim relation(10), Surrender(2), Defendant feeble-minded?(3), Defendant insanity?(3)
Drug	Drug type(14), Drug crime type(7), Defendant role(6), Narcotic handling license(6), Drug quantity(6), Profit(12), Surrender(2), Defendant feeble-minded?(3), Defendant insanity?(3)
Murder	Motivation(7), Intent to kill?(3), Self-defense?(3), Assault method(9), Injury?(3), Number of victims(3), Defendant-victim relation(10), Surrender(2), Defendant feeble-minded?(3), Defendant insanity?(3), Reason not reaching consummation(4), Reached consummation?(3)

Table 7: Factors for defining Stricter relevance. Each factor is presented with the number of options in parentheses. Question mark(?) indicates that the factor represents a *claim* defendant makes in a court, which always has three options (not mention, claimed but not taken, claimed and taken). As some factors only apply to certain standard groups (*e.g.* Traffic accident type(6) only applies to traffic crimes involving accidents and not crimes like Driving Under the Influence (without any traffic accident)) and not all combinations are possible (*e.g.* Killing Ascendant (killing one's own or any lineal ascendant of one's spouse) cases can only take two options (*parent, other family members*) out of 10 options (*partners, friend, ...*) provided for the Defendant-victim relation factor), the total number of stricter groups is a magnitude smaller compared to all option numbers multiplied.

Table 8: List of query case types for Sexual Crime

Law	Crime
형법 제297조	강간
형법 제297조의2	유사강간
형법 제298조	강제추행
형법 제299조, 제297조	준강간
형법 제299조, 제297조의2	준유사강간
형법 제299조, 제298조	준강제추행
형법 제300조, 제297조	강간미수
형법 제300조, 제297조의2	유사강간미수
형법 제300조, 제298조	강제추행미수
형법 제300조, 제297조, 제299조	준강간미수
형법 제300조, 제297조의2, 제299조	준유사강간미수
형법 제300조, 제298조, 제299조	준강제추행미수
형법 제301조, 제297조	강간상해치상
형법 제301조, 제297조의2	유사강간상해치상
형법 제301조, 제298조	강제추행상해치상
형법 제301조, 제297조, 제299조	준강간상해치상
형법 제301조, 제297조의2, 제299조	준유사강간상해치상
형법 제301조, 제298조, 제299조	준강제추행상해치상
형법 제301조, 제300조, 제297조	강간미수상해치상
형법 제301조, 제300조, 제297조의2	유사강간미수상해치상
형법 제301조, 제300조, 제298조	강제추행미수상해치상
형법 제301조, 제300조, 제297조, 제299조	준강간미수상해치상
형법 제301조, 제300조, 제297조의2, 제299조	준유사강간미수상해치상
형법 제301조, 제300조, 제298조, 제299조	준강제추행미수상해치상
형법 제301조의2, 제297조	강간살인치사
형법 제301조의2, 제297조의2	유사강간살인치사
형법 제301조의2, 제298조	강제추행살인치사
형법 제301조의2, 제297조, 제299조	준강간살인치사
형법 제301조의2, 제297조의2, 제299조	준유사강간살인치사
형법 제301조의2, 제298조, 제299조	준강제추행살인치사
형법 제301조의2, 제300조, 제297조	강간미수살인치사
형법 제301조의2, 제300조, 제297조의2	유사강간미수살인치사
형법 제301조의2, 제300조, 제298조	강제추행미수살인치사
형법 제301조의2, 제300조, 제297조, 제299조	준강간미수살인치사
형법 제301조의2, 제300조, 제297조의2, 제299조	준유사강간미수살인치사
형법 제301조의2, 제300조, 제298조, 제299조	준강제추행미수살인치사
형법 제302조	미성년자간음 / 심신미약자간음 / 미성년자추행 / 심신미약자추행
형법 제303조 제1항	피보호자간음 / 피감독자간음
형법 제305조, 제297조	미성년자의제강간
형법 제305조, 제297조의2	미성년자의제유사강간
형법 제305조, 제298조	미성년자의제강제추행
형법 제305조, 제301조, 제297조	미성년자의제강간상해치상
형법 제305조, 제301조, 제297조의2	미성년자의제유사강간상해치상
형법 제305조, 제301조, 제298조	미성년자의제강제추행상해치상
형법 제305조, 제301조의2, 제297조	미성년자의제강간살인치사
형법 제305조, 제301조의2, 제297조의2	미성년자의제유사강간살인치사
형법 제305조, 제301조의2, 제298조	미성년자의제강제추행살인치사
형법 제305조의3, 제297조	강간예비음모
형법 제305조의3, 제297조의2	유사강간예비음모
형법 제305조의3, 제297조, 제299조	준강간예비음모
형법 제305조의3, 제301조, 제297조	강간상해예비음모
형법 제305조의3, 제301조, 제298조	강제추행상해예비음모
형법 제305조의3, 제301조, 제297조, 제299조	준강간상해예비음모
형법 제305조의3, 제301조, 제297조의2, 제299조	준유사강간상해예비음모
형법 제305조의3, 제301조, 제298조, 제299조	준강제추행상해예비음모
형법 제305조의3, 제305조, 제297조	미성년자의제강간예비음모
형법 제305조의3, 제305조, 제297조의2	미성년자의제유사강간예비음모
형법 제305조의3, 제305조, 제298조	미성년자의제강제추행예비음모
형법 제305조의3, 제305조, 제301조, 제297조	미성년자의제강간상해치상예비음모
형법 제305조의3, 제305조, 제301조, 제297조의2	미성년자의제유사강간상해치상예비음모
형법 제305조의3, 제305조, 제301조, 제298조	미성년자의제강제추행상해치상예비음모
형법 제305조의3, 제305조, 제301조의2, 제297조	미성년자의제강간살인치사예비음모
형법 제305조의3, 제305조, 제301조의2, 제297조의2	미성년자의제유사강간살인치사예비음모
형법 제305조의3, 제305조, 제301조의2, 제298조	미성년자의제강제추행살인치사예비음모

Law	Crime
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제319조 제1항, 제297조	주거침입강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제319조 제1항, 제297조의2	주거침입유사강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제319조 제1항, 제298조	주거침입강제추행
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제319조 제1항, 제299조, 제297조	주거침입준강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제319조 제1항, 제299조, 제297조의2	주거침입준유사강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제319조 제1항, 제299조, 제298조	주거침입준강제추행
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제330조, 제297조	야간주거침입절도강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제330조, 제297조의2	야간주거침입절도유사강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제330조, 제298조	야간주거침입절도강제추행
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제330조, 제299조, 제297조	야간주거침입절도준강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제330조, 제299조, 제297조의2	야간주거침입절도준유사강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제330조, 제299조, 제298조	야간주거침입절도준강제추행
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제331조, 제297조	특수절도강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제331조, 제297조의2	특수절도유사강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제331조, 제298조	특수절도강제추행
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제331조, 제299조, 제297조	특수절도준강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제331조, 제299조, 제297조의2	특수절도준유사강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제331조, 제299조, 제298조	특수절도준강제추행
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제342조, 제330조, 제297조	야간주거침입절도미수강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제342조, 제330조, 제297조의2	야간주거침입절도미수유사강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제342조, 제330조, 제298조	야간주거침입절도미수강제추행
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제342조, 제330조, 제299조, 제297조	야간주거침입절도미수준강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제342조, 제330조, 제299조, 제297조의2	야간주거침입절도미수준유사강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제342조, 제330조, 제299조, 제298조	야간주거침입절도미수준강제추행
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제342조, 제331조, 제297조	특수절도미수강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제342조, 제331조, 제297조의2	특수절도미수유사강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제342조, 제331조, 제298조	특수절도미수강제추행
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제342조, 제331조, 제299조, 제297조	특수절도미수준강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제342조, 제331조, 제299조, 제297조의2	특수절도미수준유사강간
성폭력범죄의처벌등에관한특례법 제3조 제1항, 형법 제342조, 제331조, 제299조, 제298조	특수절도미수준강제추행
성폭력범죄의처벌등에관한특례법 제3조 제2항, 형법 제319조 제1항, 제297조	특수강도강간
성폭력범죄의처벌등에관한특례법 제3조 제2항, 형법 제334조, 제297조의2	특수강도유사강간
성폭력범죄의처벌등에관한특례법 제3조 제2항, 형법 제334조, 제298조	특수강도강제추행
성폭력범죄의처벌등에관한특례법 제3조 제2항, 형법 제334조, 제299조, 제297조	특수강도준강간
성폭력범죄의처벌등에관한특례법 제3조 제2항, 형법 제334조, 제299조, 제298조	특수강도준강제추행
성폭력범죄의처벌등에관한특례법 제3조 제2항, 형법 제342조, 제334조, 제297조	특수강도미수강간
성폭력범죄의처벌등에관한특례법 제3조 제2항, 형법 제342조, 제334조, 제297조의2	특수강도미수유사강간
성폭력범죄의처벌등에관한특례법 제3조 제2항, 형법 제342조, 제334조, 제298조	특수강도미수강제추행
성폭력범죄의처벌등에관한특례법 제3조 제2항, 형법 제342조, 제334조, 제299조, 제297조	특수강도미수준강간
성폭력범죄의처벌등에관한특례법 제3조 제2항, 형법 제342조, 제334조, 제299조, 제297조의2	특수강도미수준유사강간
성폭력범죄의처벌등에관한특례법 제3조 제2항, 형법 제342조, 제334조, 제299조, 제298조	특수강도미수준강제추행

Law	Crime
성폭력범죄의처벌등에관한특례법 제4조 제1항	성폭력범죄의처벌등에관한특례법위반(특수강간)
성폭력범죄의처벌등에관한특례법 제4조 제2항	성폭력범죄의처벌등에관한특례법위반(특수강제추행)
성폭력범죄의처벌등에관한특례법 제4조 제3항	성폭력범죄의처벌등에관한특례법위반(특수준강간)
성폭력범죄의처벌등에관한특례법 제4조 제3항	성폭력범죄의처벌등에관한특례법위반(특수준강제추행)
성폭력범죄의처벌등에관한특례법 제5조 제1항	성폭력범죄의처벌등에관한특례법위반(친족관계에의한강간)
성폭력범죄의처벌등에관한특례법 제5조 제2항	성폭력범죄의처벌등에관한특례법위반(친족관계에의한강제추행)
성폭력범죄의처벌등에관한특례법 제5조 제3항	성폭력범죄의처벌등에관한특례법위반(친족관계에의한준강간)
성폭력범죄의처벌등에관한특례법 제5조 제3항	성폭력범죄의처벌등에관한특례법위반(친족관계에의한준강제추행)
성폭력범죄의처벌등에관한특례법 제6조 제1항	성폭력범죄의처벌등에관한특례법위반(장애인강간)
성폭력범죄의처벌등에관한특례법 제6조 제2항	성폭력범죄의처벌등에관한특례법위반(장애인유사성행위)
성폭력범죄의처벌등에관한특례법 제6조 제3항	성폭력범죄의처벌등에관한특례법위반(장애인강제추행)
성폭력범죄의처벌등에관한특례법 제6조 제4항	성폭력범죄의처벌등에관한특례법위반(장애인준강간)
성폭력범죄의처벌등에관한특례법 제6조 제4항	성폭력범죄의처벌등에관한특례법위반(장애인준유사성행위)
성폭력범죄의처벌등에관한특례법 제6조 제4항	성폭력범죄의처벌등에관한특례법위반(장애인준강제추행)
성폭력범죄의처벌등에관한특례법 제6조 제5항	성폭력범죄의처벌등에관한특례법위반(장애인위계등간음)
성폭력범죄의처벌등에관한특례법 제6조 제6항	성폭력범죄의처벌등에관한특례법위반(장애인위계등추행)
성폭력범죄의처벌등에관한특례법 제6조 제7항	성폭력범죄의처벌등에관한특례법위반(장애인피보호자간등)
성폭력범죄의처벌등에관한특례법 제7조 제1항	성폭력범죄의처벌등에관한특례법위반(13세미만미성년자강간)
성폭력범죄의처벌등에관한특례법 제7조 제2항	성폭력범죄의처벌등에관한특례법위반(13세미만미성년자유사성행위)
성폭력범죄의처벌등에관한특례법 제7조 제3항	성폭력범죄의처벌등에관한특례법위반(13세미만미성년자강제추행)
성폭력범죄의처벌등에관한특례법 제7조 제4항	성폭력범죄의처벌등에관한특례법위반(13세미만미성년자준강간)
성폭력범죄의처벌등에관한특례법 제7조 제4항	성폭력범죄의처벌등에관한특례법위반(13세미만미성년자준유사성행위)
성폭력범죄의처벌등에관한특례법 제7조 제4항	성폭력범죄의처벌등에관한특례법위반(13세미만미성년자준강제추행)
성폭력범죄의처벌등에관한특례법 제7조 제5항	성폭력범죄의처벌등에관한특례법위반(13세미만미성년자위계등간음)
성폭력범죄의처벌등에관한특례법 제7조 제5항	성폭력범죄의처벌등에관한특례법위반(13세미만미성년자위계등유사성행위)
성폭력범죄의처벌등에관한특례법 제7조 제5항	성폭력범죄의처벌등에관한특례법위반(13세미만미성년자위계등추행)
성폭력범죄의처벌등에관한특례법 제8조	성폭력범죄의처벌등에관한특례법위반(강간등상해) / 성폭력범죄의처벌등에관한특례법위반(강간등치상)
성폭력범죄의처벌등에관한특례법 제9조	성폭력범죄의처벌등에관한특례법위반(강간등살인) / 성폭력범죄의처벌등에관한특례법위반(강간등치사)
성폭력범죄의처벌등에관한특례법 제10조	성폭력범죄의처벌등에관한특례법위반(업무상위력등에의한추행)
성폭력범죄의처벌등에관한특례법 제11조	성폭력범죄의처벌등에관한특례법위반(공중밀집장소에서의추행)
성폭력범죄의처벌등에관한특례법 제12조	성폭력범죄의처벌등에관한특례법위반(성적목적다중이용장소침입)
성폭력범죄의처벌등에관한특례법 제13조	성폭력범죄의처벌등에관한특례법위반(통신매체이용음란)
성폭력범죄의처벌등에관한특례법 제14조 제1항	성폭력범죄의처벌등에관한특례법위반(카메라등이용촬영)
성폭력범죄의처벌등에관한특례법 제14조 제2항	성폭력범죄의처벌등에관한특례법위반(카메라등이용촬영물반포등)
성폭력범죄의처벌등에관한특례법 제14조 제3항	성폭력범죄의처벌등에관한특례법위반(영리목적카메라등이용촬영물반포등)
성폭력범죄의처벌등에관한특례법 제14조 제4항	성폭력범죄의처벌등에관한특례법위반(카메라등이용촬영물소지등)
성폭력범죄의처벌등에관한특례법 제14조 제5항	성폭력범죄의처벌등에관한특례법위반(상습카메라등이용촬영·반포등)
성폭력범죄의처벌등에관한특례법 제14조의2 제1항	성폭력범죄의처벌등에관한특례법위반(허위영상물편집등)
성폭력범죄의처벌등에관한특례법 제14조의2 제2항	성폭력범죄의처벌등에관한특례법위반(허위영상물반포등)
성폭력범죄의처벌등에관한특례법 제14조의2 제3항	성폭력범죄의처벌등에관한특례법위반(영리목적허위영상물반포등)
성폭력범죄의처벌등에관한특례법 제14조의3 제1항	성폭력범죄의처벌등에관한특례법위반(촬영물등이용협박)
성폭력범죄의처벌등에관한특례법 제14조의3 제2항	성폭력범죄의처벌등에관한특례법위반(촬영물등이용강요)
성폭력범죄의처벌등에관한특례법 제14조의3 제3항, 제14조의3 제1항, 제14조 제1항	성폭력범죄의처벌등에관한특례법위반(상습촬영물등이용협박)
성폭력범죄의처벌등에관한특례법 제50조	성폭력범죄의처벌등에관한특례법위반(비밀준수등)
성매매알선등행위의처벌에관한법률 제19조	성매매알선등행위의처벌에관한법률위반(성매매알선등)
성매매알선등행위의처벌에관한법률 제20조	성매매알선등행위의처벌에관한법률위반(성매매광고)
성매매알선등행위의처벌에관한법률 제21조 제1항	성매매알선등행위의처벌에관한법률위반(성매매)
아동·청소년의성보호에관한법률 제7조 제1항	아동·청소년의성보호에관한법률위반(강간)
아동·청소년의성보호에관한법률 제7조 제2항	아동·청소년의성보호에관한법률위반(유사성행위)
아동·청소년의성보호에관한법률 제7조 제3항	아동·청소년의성보호에관한법률위반(강제추행)
아동·청소년의성보호에관한법률 제7조 제4항	아동·청소년의성보호에관한법률위반(준강간)
아동·청소년의성보호에관한법률 제7조 제4항	아동·청소년의성보호에관한법률위반(준유사성행위)
아동·청소년의성보호에관한법률 제7조 제4항	아동·청소년의성보호에관한법률위반(준강제추행)
아동·청소년의성보호에관한법률 제7조 제2항	아동·청소년의성보호에관한법률위반(위계등간음)
아동·청소년의성보호에관한법률 제7조 제2항	아동·청소년의성보호에관한법률위반(위계등유사성행위)
아동·청소년의성보호에관한법률 제7조 제2항	아동·청소년의성보호에관한법률위반(위계등추행)
아동·청소년의성보호에관한법률 제7조의2, 제7조 제1항	아동·청소년의성보호에관한법률위반(강간예비음모)
아동·청소년의성보호에관한법률 제7조의2, 제7조 제2항	아동·청소년의성보호에관한법률위반(유사성행위예비음모)
아동·청소년의성보호에관한법률 제7조의2, 제7조 제3항	아동·청소년의성보호에관한법률위반(강제추행예비음모)
아동·청소년의성보호에관한법률 제7조의2, 제7조 제4항	아동·청소년의성보호에관한법률위반(준강간예비음모)
아동·청소년의성보호에관한법률 제7조의2, 제7조 제4항	아동·청소년의성보호에관한법률위반(준유사성행위예비음모)
아동·청소년의성보호에관한법률 제7조의2, 제7조 제4항	아동·청소년의성보호에관한법률위반(준강제추행예비음모)
아동·청소년의성보호에관한법률 제7조의2, 제7조 제2항	아동·청소년의성보호에관한법률위반(위계등간음예비음모)

Law	Crime
아동·청소년의성보호에관한법률 제7조의2, 제7조 제2항	아동·청소년의성보호에관한법률위반(위계등유사성행위예비음모)
아동·청소년의성보호에관한법률 제7조의2, 제7조 제2항	아동·청소년의성보호에관한법률위반(위계등추행예비음모)
아동·청소년의성보호에관한법률 제8조 제1항	아동·청소년의성보호에관한법률위반(장애인간음)
아동·청소년의성보호에관한법률 제8조 제2항	아동·청소년의성보호에관한법률위반(장애인추행)
아동·청소년의성보호에관한법률 제8조의2 제1항	아동·청소년의성보호에관한법률위반(16세미만아동·청소년간음)
아동·청소년의성보호에관한법률 제8조의2 제2항	아동·청소년의성보호에관한법률위반(16세미만아동·청소년추행)
아동·청소년의성보호에관한법률 제9조	아동·청소년의성보호에관한법률위반(강간등상해) / 아동·청소년의성보호에관한법률위반(강간등치상)
아동·청소년의성보호에관한법률 제10조	아동·청소년의성보호에관한법률위반(강간등살인) / 아동·청소년의성보호에관한법률위반(강간등치사)
아동·청소년의성보호에관한법률 제11조 제1항	아동·청소년의성보호에관한법률위반(성착취물제작등)
아동·청소년의성보호에관한법률 제11조 제2항	아동·청소년의성보호에관한법률위반(영리목적성착취물판매등)
아동·청소년의성보호에관한법률 제11조 제3항	아동·청소년의성보호에관한법률위반(성착취물배포등)
아동·청소년의성보호에관한법률 제11조 제5항	아동·청소년의성보호에관한법률위반(성착취물소지등)
아동·청소년의성보호에관한법률 제11조 제7항	아동·청소년의성보호에관한법률위반(상습성착취물제작·배포등)
아동·청소년의성보호에관한법률 제13조 제1항	아동·청소년의성보호에관한법률위반(성매수등)
아동·청소년의성보호에관한법률 제15조의2	아동·청소년의성보호에관한법률위반(성착취목적대화등)
스토킹범죄의 처벌 등에 관한 법률 제18조 제1항	스토킹범죄의처벌등에관한법률위반
스토킹범죄의 처벌 등에 관한 법률 제20조	스토킹범죄의처벌등에관한법률위반

Table 9: List of query case types for Fraud

Law	Crime
형법 제347조	사기
형법 제347조, 제351조	상습사기
형법 제347조, 제352조	사기미수
형법 제347조, 제351조, 제352조	상습사기미수
형법 제347조의2	컴퓨터등사용사기
형법 제347조의2, 351조	상습컴퓨터등사용사기
형법 제347조의2, 352조	컴퓨터등사용사기미수
형법 제347조의2, 제351조, 제352조	상습컴퓨터등사용사기미수
형법 제348조	준사기
형법 제348조, 제351조	상습준사기
형법 제348조, 제352조	준사기미수
형법 제348조, 제351조, 제352조	상습준사기미수
형법 제348조의2	편의시설부정이용
형법 제348조의2, 제351조	상습편의시설부정이용
형법 제348조의2, 제352조	편의시설부정이용미수
형법 제348조의2, 제351조, 제352조	상습편의시설부정이용미수
형법 제350조	공갈
형법 제350조, 제351조	상습공갈
형법 제350조, 제352조	공갈미수
형법 제350조, 제351조, 제352조	상습공갈미수
형법 제350조의2	특수공갈
형법 제350조의2, 제351조	상습특수공갈
형법 제350조의2, 제352조	특수공갈미수
형법 제350조의2, 제351조, 제352조	상습특수공갈미수
폭력행위등처벌에관한법률 제2조 제2항, 형법 제350조	폭력행위등처벌에관한법률위반(공갈)
폭력행위등처벌에관한법률 제2조 제3항, 형법 제350조	폭력행위등처벌에관한법률위반(공갈재범)
특정경제범죄가중처벌등에관한법률 제3조, 형법 제347조, 제347조의2, 제351조	특정경제범죄가중처벌등에관한법률위반(사기)
특정경제범죄가중처벌등에관한법률 제3조, 형법 제350조, 제350조의2, 제351조	특정경제범죄가중처벌등에관한법률위반(공갈)
여신전문금융업법 제70조	여신전문금융업법위반

Table 10: List of query case types for Theft and Robbery

Law	Crime
형법 제329조	절도
형법 제329조, 제342조	절도미수
형법 제329조, 제332조	상습절도
형법 제329조, 제332조, 제342조	상습절도미수
형법 제330조	야간주거침입절도
형법 제330조, 제342조	야간주거침입절도미수
형법 제330조, 제332조	상습야간주거침입절도
형법 제330조, 제332조, 제342조	상습야간주거침입절도미수
형법 제330조	야간건조물침입절도
형법 제330조, 제342조	야간건조물침입절도미수
형법 제330조, 제332조	상습야간건조물침입절도
형법 제330조, 제332조, 제342조	상습야간건조물침입절도미수
형법 제330조	야간선박침입절도
형법 제330조, 제342조	야간선박침입절도미수
형법 제330조, 제332조	상습야간선박침입절도
형법 제330조, 제332조, 제342조	상습야간선박침입절도미수
형법 제330조	야간항공기침입절도
형법 제330조, 제342조	야간항공기침입절도미수
형법 제330조, 제332조	상습야간항공기침입절도
형법 제330조, 제332조, 제342조	상습야간항공기침입절도미수
형법 제330조	야간방실침입절도
형법 제330조, 제342조	야간방실침입절도미수
형법 제330조, 제332조	상습야간방실침입절도
형법 제330조, 제332조, 제342조	상습야간방실침입절도미수
형법 제331조	특수절도
형법 제331조, 제342조	특수절도미수
형법 제331조, 제332조	상습특수절도
형법 제331조, 제332조, 제342조	상습특수절도미수
형법 제333조	강도
형법 제333조, 제342조	강도미수
형법 제333조, 제341조	상습강도
형법 제333조, 제341조, 제342조	상습강도미수
형법 제334조	특수강도
형법 제334조, 제342조	특수강도미수
형법 제334조, 제341조	상습특수강도
형법 제334조, 제341조, 제342조	상습특수강도미수
형법 제335조	준강도
형법 제335조, 제342조	준강도미수
형법 제335조	준특수강도
형법 제335조, 제342조	준특수강도미수
형법 제337조	강도상해
형법 제337조, 제342조	강도상해미수
형법 제337조	강도치상
형법 제337조, 제342조	강도치상미수
형법 제338조	강도살인
형법 제338조, 제342조	강도살인미수
형법 제338조	강도치사
형법 제338조, 제342조	강도치사미수
형법 제339조	강도강간
형법 제339조, 제342조	강도강간미수
형법 제343조	강도예비
형법 제343조	강도음모

Table 11: List of query case types for Embezzlement

Law	Crime
형법 제355조 제1항	횡령
형법 제355조 제1항, 제359조	횡령미수
형법 제355조 제2항	배임
형법 제355조 제2항, 제359조	배임미수
형법 제356조	업무상횡령
형법 제356조, 제359조	업무상횡령미수
형법 제356조	업무상배임
형법 제356조, 제359조	업무상배임미수
형법 제357조 제1항	배임수재
형법 제357조 제1항, 제359조	배임수재미수
형법 제357조 제2항	배임증재
형법 제357조 제2항, 제359조	배임증재미수
형법 제360조 제1항	점유이탈물횡령
형법 제360조 제1항, 제359조	점유이탈물횡령미수
특정경제범죄가중처벌등에관한법률 제3조, 형법 제355조 제1항, 제356조	특정경제범죄가중처벌등에관한법률위반(횡령)
특정경제범죄가중처벌등에관한법률 제3조, 형법 제355조 제2항, 제356조	특정경제범죄가중처벌등에관한법률위반(배임)

Table 12: List of query case types for Destruction

Law	Crime
형법 제366조	재물손괴
형법 제369조, 제366조	특수재물손괴
형법 제369조, 제371조	특수재물손괴미수
형법 제371조, 제366조	재물손괴미수
형법 제371조, 제369조, 제366조	특수재물손괴미수
폭력행위등처벌에관한법률 제2조 제2항, 형법 제366조	폭력행위등처벌에관한법률위반(재물손괴등)
폭력행위등처벌에관한법률 제2조 제3항, 형법 제366조	폭력행위등처벌에관한법률위반(재물손괴등재범)

Table 13: List of query case types for Murder

Law	Crime
형법 제250조 제1항	살인
형법 제254조, 제250조 제1항	살인미수

Table 14: List of query case types for Injury and Violence

Law	Crime
형법 제257조 제1항	상해
형법 제257조 제1항, 제264조	상습상해
형법 제257조 제3항, 제1항	상해미수
형법 제258조 제1항/제258조 제2항	중상해
형법 제258조, 제264조	상습중상해
형법 제258조의2 제1항, 제257조 제1항	특수상해
형법 제258조의2, 제264조	상습특수상해
형법 제258조의2 제3항, 제1항, 제257조 제1항	특수상해미수
형법 제258조의2, 제264조	상습특수상해미수
형법 제259조 제1항	상해치사
형법 제260조 제1항	폭행
형법 제260조 제1항, 제264조	상습폭행
형법 제261조, 제260조 제1항	특수폭행
형법 제261조, 제264조	상습특수폭행
형법 제262조, 제260조 제1항, (제259조, 제257조 제1항)	폭행치사
형법 제262조, 제260조 제1항, (제259조, 제257조 제1항)	폭행치상
형법 제262조, 제261조, (제259조, 제257조 제1항)	특수폭행치사
형법 제262조, 제261조, (제259조, 제257조 제1항)	특수폭행치상
폭력행위등처벌에관한법률 제2조 제2항	폭력행위등처벌에관한법률위반(폭행)
폭력행위등처벌에관한법률 제2조 제2항	폭력행위등처벌에관한법률위반(존속폭행)
폭력행위등처벌에관한법률 제2조 제2항	폭력행위등처벌에관한법률위반(상해)
폭력행위등처벌에관한법률 제2조 제2항	폭력행위등처벌에관한법률위반(존속상해)
폭력행위등처벌에관한법률 제2조 제3항	폭력행위등처벌에관한법률위반(폭행재범)

Law	Crime
폭력행위등처벌에관한법률 제2조 제3항	폭력행위등처벌에관한법률위반(존속폭행재범)
폭력행위등처벌에관한법률 제2조 제3항	폭력행위등처벌에관한법률위반(상해재범)
폭력행위등처벌에관한법률 제2조 제3항	폭력행위등처벌에관한법률위반(존속상해재범)
폭력행위등처벌에관한법률 제4조 제1항	폭력행위등처벌에관한법률위반(단체등의구성·활동)
형법 제262조, 제260조 제2항	존속폭행치상
형법 제262조, 제260조 제2항	존속폭행치사
형법 제262조, 제260조 제2항	존속상해치상
형법 제262조, 제260조 제2항	존속상해치사
형법 제260조 제2항, 제264조	상습존속폭행
형법 제260조 제2항, 제264조	상습존속상해
형법 제261조, 제260조 제2항	특수존속폭행
형법 제261조, 제260조 제2항	특수존속상해
형법 제260조 제2항	존속폭행
형법 제260조 제2항	존속상해
형법 제258조 제3항	존속중상해

Table 15: List of query case types for Negligent homicide and injury

Law	Crime
형법 제266조	과실치상
형법 제267조	과실치사
형법 제268조	업무상과실치사
형법 제268조	중과실치사
형법 제268조	업무상과실치상
형법 제268조	중과실치상

Table 16: List of query case types for Arrest and Detention

Law	Crime
형법 제276조 제1항	체포
형법 제276조 제1항	감금
형법 제277조 제1항	중체포
형법 제277조 제1항	중감금
형법 제276조 제1항, 제280조	체포미수
형법 제276조 제1항, 제280조	감금미수
형법 제277조 제1항, 제280조	중체포미수
형법 제277조 제1항, 제280조	중감금미수
형법 제278조	특수체포
형법 제278조	특수감금
형법 제278조	특수중체포
형법 제278조	특수중감금
형법 제278조, 제280조	특수체포미수
형법 제278조, 제280조	특수감금미수
형법 제278조, 제280조	특수중체포미수
형법 제278조, 제280조	특수중감금미수
형법 제279조	상습체포
형법 제279조	상습감금
형법 제279조	상습중체포
형법 제279조	상습중감금
형법 제279조, 제280조	상습체포미수
형법 제279조, 제280조	상습감금미수
형법 제279조, 제280조	상습중체포미수
형법 제279조, 제280조	상습중감금미수
형법 제281조 제1항	체포치상
형법 제281조 제1항	감금치상
형법 제281조 제1항	중체포치상
형법 제281조 제1항	중감금치상
형법 제281조 제1항	체포치사
형법 제281조 제1항	감금치사
형법 제281조 제1항	중체포치사
형법 제281조 제1항	중감금치사
형법 제281조 제1항	특수존속체포
형법 제281조 제1항	상습존속체포

Law	Crime
형법 제281조 제1항	특수존속체포미수
형법 제281조 제1항	상습존속체포미수
형법 제281조 제1항	특수존속감금
형법 제281조 제1항	상습존속감금
형법 제281조 제1항	특수존속감금미수
형법 제281조 제1항	상습존속감금미수
형법 제281조 제1항	특수중존속체포
형법 제281조 제1항	상습중존속체포
형법 제281조 제1항	특수중존속체포미수
형법 제281조 제1항	상습중존속체포미수
형법 제281조 제1항	특수중존속감금
형법 제281조 제1항	상습중존속감금
형법 제281조 제1항	특수중존속감금미수
형법 제281조 제1항	상습중존속감금미수
형법 제277조 제2항	중존속체포
형법 제277조 제2항	중존속감금
형법 제276조 제2항	존속체포
형법 제276조 제2항	존속감금
폭력행위등처벌에관한법률 제2조 제2항, 형법 제276조 제1항	폭력행위등처벌에관한법률위반(체포)
폭력행위등처벌에관한법률 제2조 제2항, 형법 제276조 제1항	폭력행위등처벌에관한법률위반(감금)
폭력행위등처벌에관한법률 제2조 제2항, 형법 제276조 제2항	폭력행위등처벌에관한법률위반(존속체포)
폭력행위등처벌에관한법률 제2조 제2항, 형법 제276조 제2항	폭력행위등처벌에관한법률위반(존속감금)
폭력행위등처벌에관한법률 제2조 제3항, 형법 제276조 제1항	폭력행위등처벌에관한법률위반(체포재범)
폭력행위등처벌에관한법률 제2조 제3항, 형법 제276조 제1항	폭력행위등처벌에관한법률위반(감금재범)
폭력행위등처벌에관한법률 제2조 제3항, 형법 제276조 제2항	폭력행위등처벌에관한법률위반(존속체포재범)
폭력행위등처벌에관한법률 제2조 제3항, 형법 제276조 제2항	폭력행위등처벌에관한법률위반(존속감금재범)

Table 17: List of query case types for Threat

Law	Crime
형법 제283조 제1항	협박
형법 제284조	특수협박
형법 제285조	상습협박
형법 제285조, 제284조	상습특수협박
형법 제286조	협박미수
형법 제284조	특수협박미수
형법 제285조	상습협박미수
형법 제285조, 제284조	상습특수협박미수
형법 제283조 제2항	존속협박
형법 제284조, 제283조 제2항	특수존속협박
형법 제285조, 제283조 제2항	상습존속협박
폭력행위등처벌에관한법률 제2조 제2항, 형법 제283조 제1항	폭력행위등처벌에관한법률위반(협박)
폭력행위등처벌에관한법률 제2조 제2항, 형법 제283조 제2항	폭력행위등처벌에관한법률위반(존속협박)
폭력행위등처벌에관한법률 제2조 제3항, 형법 제283조 제1항	폭력행위등처벌에관한법률위반(협박재범)
폭력행위등처벌에관한법률 제2조 제3항, 형법 제283조 제2항	폭력행위등처벌에관한법률위반(존속협박재범)

Table 18: List of query case types for Defamation and Insult

Law	Crime
형법 제307조 제1항	사실적시명예훼손
형법 제307조 제2항	허위적시명예훼손
형법 제309조, (제307조 제1항, 제307조 제2항)	출판물에의한명예훼손
형법 제309조, (제307조 제1항, 제307조 제2항)	라디오에의한명예훼손
형법 제311조	모욕
정보통신망이용촉진및정보보호등에관한법률 제70조 제1항	정보통신망이용촉진및정보보호등에관한법률위반(사실적시명예훼손)
정보통신망이용촉진및정보보호등에관한법률 제70조 제2항	정보통신망이용촉진및정보보호등에관한법률위반(허위적시명예훼손)
정보통신망이용촉진및정보보호등에관한법률 제71조 제1항	정보통신망이용촉진및정보보호등에관한법률위반(정보통신망침해등)
정보통신망이용촉진및정보보호등에관한법률 제74조 제1항 제2호	정보통신망이용촉진및정보보호등에관한법률위반(음란물유포)

Table 19: List of query case types for Obstruction of business

Law	Crime
형법 제314조 제1항	업무방해
형법 제314조 제1항	공무집행방해
형법 제314조 제1항	위계에의한공무집행방해
형법 제314조 제1항	공용서류손상
형법 제314조 제1항	공용서류은닉
형법 제314조 제1항	공용서류무효
형법 제314조 제1항	공용물건손상
형법 제314조 제1항	공용물건은닉
형법 제314조 제1항	공용물건무효
형법 제314조 제1항	공용전자기록등손상
형법 제314조 제1항	공용전자기록등은닉
형법 제314조 제1항	공용전자기록등무효
형법 제314조 제1항	특수공무집행방해치사
형법 제314조 제1항	특수공무집행방해치상
형법 제314조 제1항	특수공무집행방해
형법 제314조 제1항	특수공용물건손상

Table 20: List of query case types for Home invasion

Law	Crime
형법 제319조 제1항	주거침입
형법 제319조 제1항	건조물침입
형법 제319조 제1항	방실침입
형법 제319조 제2항	퇴거불응
형법 제320조	특수주거침입
형법 제320조	특수건조물침입
형법 제320조	특수방실침입
형법 제320조	특수퇴거불응
형법 제322조	주거침입미수
형법 제322조	건조물침입미수
형법 제322조	방실침입미수
형법 제322조	퇴거불응미수
형법 제320조, 제322조	특수주거침입미수
형법 제320조, 제322조	특수건조물침입미수
형법 제320조, 제322조	특수방실침입미수
형법 제320조, 제322조	특수퇴거불응미수
폭력행위등처벌에관한법률 제2조 제2항, 형법 제319조 제1항	폭력행위등처벌에관한법률위반(주거침입)
폭력행위등처벌에관한법률 제2조 제2항, 형법 제319조 제2항	폭력행위등처벌에관한법률위반(퇴거불응)
폭력행위등처벌에관한법률 제2조 제3항, 형법 제319조 제1항	폭력행위등처벌에관한법률위반(주거침입재범)
폭력행위등처벌에관한법률 제2조 제3항, 형법 제319조 제2항	폭력행위등처벌에관한법률위반(퇴거불응재범)

Table 21: List of query case types for Obstruction of right

Law	Crime
형법 제323조	권리행사방해
형법 제324조 제1항	강요
형법 제324조 제2항	특수강요
형법 제327조	강제집행면탈
폭력행위등처벌에관한법률 제2조 제2항, 형법 제324조 제1항	폭력행위등처벌에관한법률위반(강요)
폭력행위등처벌에관한법률 제2조 제3항, 형법 제324조 제1항	폭력행위등처벌에관한법률위반(강요재범)

Table 22: List of query case types for Drug

Law	Crime
마약류관리에관한법률 제2조 제2호	마약류관리에관한법률위반(마약)
마약류관리에관한법률 제5조의2 제5항	마약류관리에관한법률위반임시마약류마약
마약류관리에관한법률 제2조 제3호	마약류관리에관한법률위반(향정)
마약류관리에관한법률 제5조의2 제5항	마약류관리에관한법률위반임시마약류향정
마약류관리에관한법률 제2조 제4호	마약류관리에관한법률위반(대마)
마약류관리에관한법률 제5조의2 제5항	마약류관리에관한법률위반임시마약류대마

Table 23: List of query case types for Gambling

Law	Crime
형법 제246조 제1항	도박
형법 제246조 제2항	상습도박
형법 제247조	도박장소개설
형법 제247조	도박공간개설
국민체육진흥법 제47조 제2호	국민체육진흥법위반(도박개장등)
국민체육진흥법 제48조 제3호, 제26조 제1항	국민체육진흥법위반(도박등)
국민체육진흥법 제48조 제4호, 제26조 제2항 제1호	국민체육진흥법위반(도박개장등)
게임산업진흥에관한법률	게임산업진흥에관한법률위반

Table 24: List of query case types for Other criminal offenses

Law	Crime
형법 제145조	도주
형법 제151조	범인은닉
형법 제151조	범인도피
형법 제152조 제1항	위증
형법 제156조	무고
형법 제164조 제1항, 제165조, 제166조, 제167조	방화
형법 제164조 제2항, 제164조 제1항	방화치상
형법 제164조 제2항, 제164조 제1항	방화치사
형법 제170조	실화
형법 제171조	업무상실화
형법 제174조	방화미수
형법 제175조	방화예비
형법 제185조	일반교통방해
형법 제186조	기차교통방해
형법 제186조	전차교통방해

Table 25: List of query case types for Traffic offenses

Law	Crime
교통사고처리특례법 제3조 제1항, 형법 제268조	교통사고처리특례법위반(치사)
교통사고처리특례법 제3조 제1항, 형법 제268조	교통사고처리특례법위반(치상)
교통사고처리특례법	교통사고처리특례법위반
특정범죄가중처벌등에관한법률 제5조의3 제1항 제1호	특정범죄가중처벌등에관한법률위반(도주치사)
특정범죄가중처벌등에관한법률 제5조의3 제1항 제2호	특정범죄가중처벌등에관한법률위반(도주치상)
특정범죄가중처벌등에관한법률 제5조의10	특정범죄가중처벌등에관한법률위반(운전자폭행등)
특정범죄가중처벌등에관한법률 제5조의11 제1항	특정범죄가중처벌등에관한법률위반(위험운전치사)
특정범죄가중처벌등에관한법률 제5조의11 제1항	특정범죄가중처벌등에관한법률위반(위험운전치상)
특정범죄가중처벌등에관한법률 제5조의13	특정범죄가중처벌등에관한법률위반(어린이보호구역치사)
특정범죄가중처벌등에관한법률 제5조의13	특정범죄가중처벌등에관한법률위반(어린이보호구역치상)
도로교통법 제43조	도로교통법위반(무면허운전)
도로교통법 제44조 제1항	도로교통법위반(음주운전)
도로교통법 제44조 제2항	도로교통법위반(음주측정거부)
도로교통법 제46조	도로교통법위반(공동위험행위)
도로교통법 제54조 제1항	도로교통법위반(사고후미조치)

Table 26: List of query case types for Child crime and School violence

Law	Crime
아동복지법 제71조 제1항 제2호, 제17조 각호	아동복지법위반(아동학대)
아동복지법	아동복지법위반(아동유기 · 방임)
아동복지법	아동복지법위반(상습아동학대)
아동복지법	아동복지법위반(상습아동유기 · 방임)
아동학대범죄의처벌등에관한특례법 제4조 제2항	아동학대범죄의처벌등에관한특례법위반(아동학대치사)
아동학대범죄의처벌등에관한특례법 제5조	아동학대범죄의처벌등에관한특례법위반(아동학대중상해)

Law	Crime
아동학대범죄의처벌등에관한특례법 제7조	아동학대범죄의처벌등에관한특례법위반(아동복지시설종사자등의아동학대)
청소년보호법	청소년보호법위반
학교폭력예방및대책에관한법률	학교폭력예방및대책에관한법률위반

Table 27: List of query case types for Corporation

Law	Crime
부정경쟁방지법 제18조 제1항	부정경쟁방지및영업비밀보호에관한법률위반(영업비밀국외누설등)
부정경쟁방지법 제18조 제2항	부정경쟁방지및영업비밀보호에관한법률위반(영업비밀누설등)
자본시장법	자본시장과금융투자업에관한법률위반

Table 28: List of query case types for Labor and Employment

Law	Crime
근로기준법 제107조, 제7조	근로기준법위반(강제근로의금지)
근로기준법 제107조, 제8조	근로기준법위반(폭행의금지)
근로기준법 제107조, 제9조	근로기준법위반(중간착취의배제)
근로기준법 제107조, 제23조 제2항	근로기준법위반(해고등의제한)
근로기준법 제107조, 제40조	근로기준법위반(취업방해의금지)
근로기준법 제109조 제1항, 제36조	근로기준법위반(금품청산)
근로기준법 제109조 제1항, 제43조	근로기준법위반(임금지급)
근로기준법 제109조 제1항, 제46조	근로기준법위반(휴업수당지급)
근로기준법 제109조 제1항, 제51조의2	근로기준법위반(가산임금지급)
근로기준법 제109조 제1항, 제52조 제2항 제2호	근로기준법위반(가산임금지급)
근로기준법 제109조 제1항, 제56조	근로기준법위반(가산임금지급)
근로기준법 제109조 제1항, 제65조	근로기준법위반(사용금지)
근로기준법 제109조 제1항, 제72조	근로기준법위반(갱내근로의금지)
근로기준법 제109조 제1항, 제76조의3 제6항	근로기준법위반(직장내괴롭힘발생시조치)
근로기준법 제110조 제1호, 제10조	근로기준법위반(공민권행사보장)
근로기준법 제110조 제1호, 제22조 제1항	근로기준법위반(강제저금의금지)
근로기준법 제110조 제1호, 제26조	근로기준법위반(해고의예고)
근로기준법 제110조 제1호, 제50조	근로기준법위반(법정근로시간)
근로기준법 제110조 제1호, 제51조의2 제2항	근로기준법위반(휴식시간제공)
근로기준법 제110조 제1호, 제52조 제2항 제1호	근로기준법위반(휴식시간제공)
근로기준법 제110조 제1호, 제53조 제1항	근로기준법위반(연장근로의제한)
근로기준법 제110조 제1호, 제53조 제2항	근로기준법위반(연장근로의제한)
근로기준법 제110조 제1호, 제53조 제4항 본문	근로기준법위반(연장근로의제한)
근로기준법 제110조 제1호, 제53조 제7항	근로기준법위반(연장근로의제한)
근로기준법 제110조 제1호, 제54조	근로기준법위반(휴식시간제공)
근로기준법 제110조 제1호, 제55조	근로기준법위반(휴일보장)
근로기준법 제110조 제1호, 제59조 제2항	근로기준법위반(휴식시간제공)
근로기준법 제110조 제1호, 제60조 제1항	근로기준법위반(연차유급휴가지급)
근로기준법 제110조 제1호, 제60조 제2항	근로기준법위반(연차유급휴가지급)
근로기준법 제110조 제1호, 제60조 제4항	근로기준법위반(연차유급휴가지급)
근로기준법 제110조 제1호, 제60조 제5항	근로기준법위반(연차유급휴가지급)
근로기준법 제110조 제1호, 제69조	근로기준법위반(미성년자근로시간)
근로기준법 제110조 제1호, 제70조 제1항	근로기준법위반(야간근로휴일근로)
근로기준법 제110조 제1호, 제70조 제2항	근로기준법위반(야간근로휴일근로)
근로기준법 제110조 제1호, 제71조	근로기준법위반(시간외근로)
근로기준법 제110조 제1호, 제74조	근로기준법위반(임산부보호)
근로기준법 제110조 제1호, 제75조	근로기준법위반(육아시간제공)
근로기준법 제110조 제1호, 제104조 제2항	근로기준법위반(불리한처우금지)
근로기준법 제110조 제2호, 제53조 제5항	근로기준법위반(휴식시간제공)
근로기준법 제114조 제1호, 제6조	근로기준법위반(균등한처우)
근로기준법 제114조 제1호, 제17조	근로기준법위반(근로조건명시)
근로기준법 제114조 제1호, 제20조	근로기준법위반(위약예정금지)
근로기준법 제114조 제1호, 제21조	근로기준법위반(전차금상계금지)
근로기준법 제114조 제1호, 제22조 제2항	근로기준법위반(강제저금금지)
근로기준법 제114조 제1호, 제67조 제1항	근로기준법위반(근로계약대리금지)
근로기준법 제114조 제1호, 제67조 제3항	근로기준법위반(근로조건서면교부)
근로기준법 제114조 제1호, 제70조 제3항	근로기준법위반(야간근로휴일근로제한)
근로기준법 제114조 제1호, 제73조	근로기준법위반(생리휴가지급)
근로기준법 제114조 제1호, 제74조 제6항	근로기준법위반(임산부의보호)

Law	Crime
근로기준법 제114조 제1호, 제94조	근로기준법위반(취업규칙작성변경)
근로기준법 제114조 제1호, 제95조	근로기준법위반(감급제재제한규정)
근로기준법 제116조 제1항, 제76조의2	근로기준법위반(직장내괴롭힘금지)
근로기준법 제116조 제2항 제2호, 제41조	근로기준법위반(근로자명부작성)
근로기준법 제116조 제2항 제2호, 제42조	근로기준법위반(계약서류보존)
근로기준법 제116조 제2항 제2호, 제48조	근로기준법위반(임금대장및임금내역서)
근로기준법 제116조 제2항 제2호, 제74조 제7항	근로기준법위반(임산부의보호)
근로기준법 제116조 제2항 제2호, 제74조 제9항	근로기준법위반(임산부의보호)
근로기준법 제116조 제2항 제2호, 제76조의3 제2항	근로기준법위반(직장내괴롭힘발생시조치)
근로기준법 제116조 제2항 제2호, 제76조의3 제4항	근로기준법위반(직장내괴롭힘발생시조치)
근로기준법 제116조 제2항 제2호, 제76조의3 제5항	근로기준법위반(직장내괴롭힘발생시조치)
근로기준법 제116조 제2항 제2호, 제76조의3 제7항	근로기준법위반(직장내괴롭힘발생시조치)
근로기준법 제116조 제2항 제2호, 제91조	근로기준법위반(서류보존)
근로기준법 제116조 제2항 제2호, 제93조	근로기준법위반(취업규칙작성신고)
근로기준법 제116조 제2항 제3호, 제51조의2 제5항	근로기준법위반(임금보전방안마련)
퇴직급여법	근로자퇴직급여보장법위반
최저임금법	최저임금법위반
노동조합법 제88조, 제41조	노동조합및노동관계조정법위반(쟁의행위제한과금지)
노동조합법 제89조, 제37조	노동조합및노동관계조정법위반(노동조합아닌자에의한쟁의행위금지)
노동조합법 제89조, 제38조	노동조합및노동관계조정법위반(노동조합의지도와책임)
노동조합법 제89조, 제42조	노동조합및노동관계조정법위반(폭력행위등의금지)
노동조합법 제89조, 제42조의2	노동조합및노동관계조정법위반(필수유지업무에대한쟁의행위제한)
노동조합법 제90조, 제44조	노동조합및노동관계조정법위반(쟁의행위기간중의임금지급요구금지)
노동조합법 제90조, 제77조	노동조합및노동관계조정법위반(긴급조정시쟁의행위중지)
노동조합법 제90조, 제81조	노동조합및노동관계조정법위반(부당노동행위)
노동조합법 제91조, 제38조	노동조합및노동관계조정법위반(노동조합의지도와책임)
노동조합법 제91조, 제41조	노동조합및노동관계조정법위반(쟁의행위제한금지)
노동조합법 제91조, 제42조	노동조합및노동관계조정법위반(폭력행위등의금지)
노동조합법 제91조, 제43조	노동조합및노동관계조정법위반(사용자의채용제한)
노동조합법 제91조, 제46조	노동조합및노동관계조정법위반(직장폐쇄)
노동조합법 제91조, 제63조	노동조합및노동관계조정법위반(중재시쟁의행위금지)
노동조합법 제92조, 제31조	노동조합및노동관계조정법위반(단체협약준수)
노동조합법 제93조, 제7조	노동조합및노동관계조정법위반(노동조합의보호요건)
노동조합법 제93조, 제21조, 제31조	노동조합및노동관계조정법위반(시정명령)
노동조합법 제96조, 제14조	노동조합및노동관계조정법위반(서류비치)
노동조합법 제96조, 제27조	노동조합및노동관계조정법위반(자료제출)
노동조합법 제96조, 제46조	노동조합및노동관계조정법위반(직장폐쇄신고)
노동조합법 제96조, 제13조, 제28조제2항, 제31조 제2항	노동조합및노동관계조정법위반(신고의무)

Table 29: List of query case types for Industrial accident and Serious accident

Law	Crime
산업안전법 제167조, 제38조 제1,2,3항, 제166조의2	산업안전보건법위반(안전조치)
산업안전법 제167조, 제39조 제1항	산업안전보건법위반(보건조치)
산업안전법 제167조, 제63조	산업안전보건법위반(안전보건조치)
산업안전법 제168조, 제38조 제1,2,3항	산업안전보건법위반(안전조치)
산업안전법 제168조, 제39조 제1항	산업안전보건법위반(보건조치)
산업안전법 제168조, 제51조	산업안전보건법위반(작업중지의무)
산업안전법 제168조, 제54조 제1항	산업안전보건법위반(사업주의조치)
산업안전법 제168조, 제117조 제1항	산업안전보건법위반(유해위험물질제조등금지)
산업안전법 제168조, 제118조 제1항	산업안전보건법위반(무허가유해위험물질제조)
산업안전법 제168조, 제122조 제1항	산업안전보건법위반(석면해체제거)
산업안전법 제168조, 제157조 제3항	산업안전보건법위반(불리한처우금지)
산업안전법 제169조, 제45조 제1항, 제46조 제5항, 제53조 제1항, 제87조 제2항, 제118조 제4항, 제119조 제4항, 제131조 제1항	산업안전보건법위반(명령위반)
산업안전법 제170조, 제41조제3항	산업안전보건법위반(불리한처우금지)
산업안전법 제170조, 제56조제3항	산업안전보건법위반(중대재해발생현장훼손)
산업안전법 제170조, 제57조제1항	산업안전보건법위반(산업재해사실은폐교사공모)
산업안전법 제170조, 제153조	산업안전보건법위반(자격증등록증대여)
산업안전법 제171조, 제125조제6항, 제132조제4항	산업안전보건법위반(조치의무)
산업안전법 제172조, 제64조 제1항 제7호, 제64조 제1항 제8호, 제64조 제2항	산업안전보건법위반(산업재해예방조치의무)
산재보험법	산업재해보상보험법위반
중대재해처벌법	중대재해처벌등에관한법률위반

Table 30: List of query case types for Medical or Food drug

Law	Crime
약사법 제93조 제1항 제1호, 제6조 제3항	약사법위반(약사면허대여)
약사법 제93조 제1항 제2호, 제20조 제1항	약사법위반(약국개설금지)
약사법 제93조 제1항 제3호, 제23조 제1항	약사법위반(무면허의약품조제)
약사법 제93조 제1항 제4호, 제31조 제1항, 제2항, 제3항, 제4항, 제9항	약사법위반(허가,신고미이행)
약사법 제93조 제1항 제5호, 제42조 제1항	약사법위반(무허가,무신고의약품수입등)
약사법 제93조 제1항 제7호, 제44조 제1항	약사법위반((약국개설자가아님에도)의약품판매등)
약사법 제93조 제1항 제10호, 제61조 제1항 제1호	약사법위반(위조의약품판매등)
약사법 제93조 제1항 제10호, 제61조 제1항 제2호	약사법위반(무허가,무신고의약품판매등)
약사법 제93조 제1항 제10호, 제61조 제2항	약사법위반(의약품유사광고등)
약사법 제93조 제1항 제10호, 제66조	약사법위반(의약외품)
약사법 제94조 제1항 제2호, 제24조 제2항	약사법위반(담합행위)
약사법 제94조 제1항 제5호의4, 제47조 제2항	약사법위반(경제적이익등제공)
약사법 제94조 제1항 제5호의4, 제47조 제6항	약사법위반(경제적이익등취득)
약사법 제94조 제1항 제8호, 제50조 제1항	약사법위반(점포이외의장소에서의약품판매)
약사법 제94조 제1항 제9호, 제62조	약사법위반(유해/위해의약품판매등)
약사법 제94조 제1항 제9호, 제66조, 제62조	약사법위반(유해/위해의약외품판매등)
약사법 제95조 제1항 제2호, 제21조 제1항	약사법위반(약국이중개설)
약사법 제95조 제1항 제8호, 제47조 제1항, 제4항, 제7항	약사법위반(유통체계확립및판매질서유지의무위반)
약사법 제95조 제1항 제10호, 제60조	약사법위반(기재금지사항)
약사법 제95조 제1항 제10호, 제68조	약사법위반(의약품과장광고등)
식품위생법 제94조 제1항 제1호, 제4조	식품위생법위반(위해식품등의판매등)
식품위생법 제94조 제1항 제3호, 제37조 제1항	식품위생법위반(무허가유흥주점)
식품위생법 제95조 제1호, 제7조 제4항	식품위생법위반(식품또는식품첨가물에관한기준및규격)
식품위생법 제95조 제1호, 제9조 제4항	식품위생법위반(기구및용기포장에관한기준및규격)
식품위생법 제95조 제2호의2, 제37조 제5항	식품위생법위반(영업허가등(허가,등록,신고))
식품위생법 제96조, 제51조, 제52조	식품위생법위반(조리사미배치,영양사미채용)
식품위생법 제97조 제1호, 제37조 제4항	식품위생법위반(무허가,무신고영업등)
식품위생법 제97조 제4호, 제36조	식품위생법위반(시설기준위반)
식품위생법 제97조 제6호, 제44조 제1항	식품위생법위반(영업자등준수사항위반)
식품위생법 제97조 제7호, 제75조 제1항	식품위생법위반(영업정지명령위반)
식품위생법 제98조 제1호, 제44조 제3항	식품위생법위반(접객행위알선)

Table 31: List of query case types for Fair trade

Law	Crime
공정거래법 제6조, 제5조 제1항 제1호, 제124조 제1항 제1호	독점규제및공정거래에관한법률위반(시장지배적지위남용행위-가격남용행위)
공정거래법 제6조, 제5조 제1항 제2호, 제124조 제1항 제1호	독점규제및공정거래에관한법률위반(시장지배적지위남용행위-출고조절행위)
공정거래법 제6조, 제5조 제1항 제3호, 제124조 제1항 제1호	독점규제및공정거래에관한법률위반(시장지배적지위남용행위-사업활동방해행위)
공정거래법 제6조, 제5조 제1항 제4호, 제124조 제1항 제1호	독점규제및공정거래에관한법률위반(시장지배적지위남용행위-진입제한행위)
공정거래법 제6조, 제5조 제1항 제5호, 제124조 제1항 제1호	독점규제및공정거래에관한법률위반(시장지배적지위남용행위-경쟁사업자배제또는소비자이익의저해행위)
공정거래법 제13조, 제9조, 제124조 제1항 제2호	독점규제및공정거래에관한법률위반(경쟁제한적기업결합행위)
공정거래법 제19조, 제124조 제1항 제5호	독점규제및공정거래에관한법률위반(상호출자제한기업집단의지주회사설립제한위반행위)
공정거래법 제20조, 제124조 제1항 제6호	독점규제및공정거래에관한법률위반(일반지주회사의금융회사주식소유제한위반행위)
공정거래법 제21조, 제124조 제1항 제7호	독점규제및공정거래에관한법률위반(상호출자금지위반행위)
공정거래법 제24조, 제124조 제1항 제8호	독점규제및공정거래에관한법률위반(채무보증제한위반행위)
공정거래법 제25조, 제124조 제1항 제3호	독점규제및공정거래에관한법률위반(금융보험사의결권제한위반행위)
공정거래법 제17조, 제18조, 제124조 제1항 제4호, 제126조 제1호, 제126조 제2호	독점규제및공정거래에관한법률위반(지주회사등의행위제한위반행위)
공정거래법 제26조, 제130조 제1항 제4호, 제37조 제1항 제7호	독점규제및공정거래에관한법률위반(대규모내부거래에대한이사회의결및공시제도위반)
공정거래법 제27조, 제130조 제1항 제4호, 제37조 제1항 제7호	독점규제및공정거래에관한법률위반(비상장회사등의중요사항공시제도위반행위)
공정거래법 제28조, 제130조 제1항 제4호, 제37조 제1항 제7호	독점규제및공정거래에관한법률위반(기업집단현황공시제도위반행위)
공정거래법 제22조, 제124조 제1항 제7호	독점규제및공정거래에관한법률위반(순환출자금지위반행위)
공정거래법 제23조, 제124조 제1항 제3호	독점규제및공정거래에관한법률위반(순환출자에대한의결권제한위반행위)
공정거래법 제40조 제1항 제1호, 제124조 제1항 제9호	독점규제및공정거래에관한법률위반(부당공동행위-가격의결정,유지,변경행위)

Law	Crime
공정거래법 제40조 제1항 제2호, 제124조 제1항 제9호	독점규제및공정거래에관한법률위반(부당공동행위-거래조건및대금지급조건설정행위)
공정거래법 제40조 제1항 제3호, 제124조 제1항 제9호	독점규제및공정거래에관한법률위반(부당공동행위-거래제한행위)
공정거래법 제40조 제1항 제4호, 제124조 제1항 제9호	독점규제및공정거래에관한법률위반(부당공동행위-시장분할행위)
공정거래법 제40조 제1항 제5호, 제124조 제1항 제9호	독점규제및공정거래에관한법률위반(부당공동행위-설비제한행위)
공정거래법 제40조 제1항 제6호, 제124조 제1항 제9호	독점규제및공정거래에관한법률위반(부당공동행위-상품의종류,규격제한행위)
공정거래법 제40조 제1항 제7호, 제124조 제1항 제9호	독점규제및공정거래에관한법률위반(부당공동행위-영업의주요부문공동관리행위)
공정거래법 제40조 제1항 제8호, 제124조 제1항 제9호	독점규제및공정거래에관한법률위반(부당공동행위-입찰담합행위)
공정거래법 제40조 제1항 제9호, 제124조 제1항 제9호	독점규제및공정거래에관한법률위반(부당공동행위-정보교환공동행위)
공정거래법 제51조 제1항 제1호, 제124조 제1항 제12호, 제53조, 제52조	독점규제및공정거래에관한법률위반(사업자단체금지행위-부당한공동행위)
공정거래법 제51조 제1항 제2호, 제53조, 제52조	독점규제및공정거래에관한법률위반(사업자단체금지행위-사업자수제한행위)
공정거래법 제51조 제1항 제3호, 제125조 제5호, 제53조, 제52조	독점규제및공정거래에관한법률위반(사업자단체금지행위-사업활동방해행위)
공정거래법 제51조 제1항 제4호, 제53조, 제52조	독점규제및공정거래에관한법률위반(사업자단체금지행위-불공정거래행위)
공정거래법 제45조 제1항 제1호	독점규제및공정거래에관한법률위반(일반불공정거래행위-거래거절행위)
공정거래법 제45조 제1항 제2호	독점규제및공정거래에관한법률위반(일반불공정거래행위-차별적취급행위)
공정거래법 제45조 제1항 제3호	독점규제및공정거래에관한법률위반(일반불공정거래행위-경쟁사업자배제행위)
공정거래법 제45조 제1항 제4호, 제125조 제4호	독점규제및공정거래에관한법률위반(일반불공정거래행위-부당한고객유인행위)
공정거래법 제45조 제1항 제5호, 제125조 제4호	독점규제및공정거래에관한법률위반(일반불공정거래행위-거래강제행위)
공정거래법 제45조 제1항 제6호, 제125조 제4호	독점규제및공정거래에관한법률위반(일반불공정거래행위-거래상지위남용행위)
공정거래법 제45조 제1항 제7호	독점규제및공정거래에관한법률위반(일반불공정거래행위-구속조건부거래행위)
공정거래법 제45조 제1항 제8호, 제125조 제4호	독점규제및공정거래에관한법률위반(일반불공정거래행위-사업활동방해행위)
공정거래법 제45조 제1항 제9호, 령 별표2 제9호 가목, 제124조 제1항 제10호	독점규제및공정거래에관한법률위반(부당지원행위-부당한자금지원행위)
공정거래법 제45조 제1항 제9호, 령 별표2 제9호 나목, 제124조 제1항 제10호	독점규제및공정거래에관한법률위반(부당지원행위-부당한자산상품등지원행위)
공정거래법 제45조 제1항 제9호, 령 별표2 제9호 다목, 제124조 제1항 제10호	독점규제및공정거래에관한법률위반(부당지원행위-부당한인력지원행위)
공정거래법 제45조 제1항 제9호, 령 별표2 제9호 라목, 제124조 제1항 제10호	독점규제및공정거래에관한법률위반(부당지원행위-부당한거래단계추가행위)
공정거래법 제48조, 제124조 제1항 제11호	독점규제및공정거래에관한법률위반(보복조치행위)
공정거래법 제47조, 제124조 제1항 제10호	독점규제및공정거래에관한법률위반(특수관계인에대한부당한이익제공행위)
공정거래법 제81조 제2항, 제124조 제1항 제13호	독점규제및공정거래에관한법률위반(조사시폭언폭행등행위)
공정거래법 제30조, 제126조 제3호	독점규제및공정거래에관한법률위반(주식소유현황신고의무위반행위)
표시광고법	표시·광고의공정화에관한법률위반

Table 32: List of query case types for Finance or Insurance

Law	Crime
특정금융거래정보의보고및이용등에관한법률	특정금융거래정보의보고및이용등에관한법률위반
금융실명거래및비밀보장에관한법률	금융실명거래및비밀보장에관한법률위반
대부업법 제3조, 제3조의2, 제9조의2, 제19조 제1항	대부업등의등록및금융이용자보호에관한법률(미등록대부업및광고행위)
대부업법 제5조의2 제5항, 제19조 제2항 제1호	대부업등의등록및금융이용자보호에관한법률(타인명의대부행위)
대부업법 제8조, 제11조 제1항, 제19조 제2항 제2호	대부업등의등록및금융이용자보호에관한법률(이자율제한행위)
대부업법 제9조의3 제2항, 제19조 제2항 제3호	대부업등의등록및금융이용자보호에관한법률(허위과장광고금지행위)
대부업법 제9조의9 제1항, 제2항, 제19조 제2항 제4호	대부업등의등록및금융이용자보호에관한법률(대부업이용자의개인정보보호처리행위)
대부업법 제9조의9 제3항, 제19조 제2항 제5호	대부업등의등록및금융이용자보호에관한법률(대부업이용자의개인정보제공보관등행위)
대부업법 제10조 제1항, 제7항, 제19조 제3항	대부업등의등록및금융이용자보호에관한법률(신용공여행위)
대부업법 제5조의2 제4항, 제19조 제4항 제1호	대부업등의등록및금융이용자보호에관한법률(상호사용행위)
대부업법 제7조 제3항, 제19조 제4항 제2호	대부업등의등록및금융이용자보호에관한법률(서류목적외사용행위)
대부업법 제9조의4 제1항, 제2항, 제19조 제4항 제4호	대부업등의등록및금융이용자보호에관한법률(불법사금융업자로부터의채권양수및추심,대부중개행위)
대부업법 제9조의4 제3항, 제19조 제4항 제5호	대부업등의등록및금융이용자보호에관한법률(불법사금융업자에의채권양도행위)
대부업법 제11조의2 제1항, 제2항, 제19조 제4항 제6호	대부업등의등록및금융이용자보호에관한법률(대부중개행위)
대부업법 제11조의2 제3항, 제19조 제4항 제7호	대부업등의등록및금융이용자보호에관한법률(중개수수료초과지급행위)

Law	Crime
대부업법 제11조의2 제6항, 제19조 제4항 제9호	대부업등의등록및금융이용자보호에관한법률(중개수수료수취행위)
보험사기방지특별법	보험사기방지특별법위반
유사수신행위의규제에관한법률	유사수신행위의규제에관한법률위반
전자금융거래법	전자금융거래법위반
범죄수익은닉의규제및처벌등에관한법률	범죄수익은닉의규제및처벌등에관한법률위반

Table 33: List of query case types for Tax or Administrative or Constitutional Law

Law	Crime
변호사법	변호사법위반
조세범처벌법	조세범처벌법위반
국가보안법	국가보안법위반
출입국관리법	출입국관리법위반
주민등록법	주민등록법위반
전기통신사업법	전기통신사업법위반
공직선거법	공직선거법위반
국민체육진흥법	국민체육진흥법위반
도시및주거환경정비법	도시및주거환경정비법위반
집회및시위에관한법률	집회및시위에관한법률위반
공인중개사법	공인중개사법위반
건설산업기본법	건설산업기본법위반
보조금관리에관한법률	보조금관리에관한법률위반
성폭력범죄의처벌및피해자보호등에관한법률	성폭력범죄의처벌및피해자보호등에관한법률위반

Table 34: List of query case types for Sexual morality

Law	Crime
형법 제245조	공연음란

Table 35: List of query case types for Misdemeanor

Law	Crime
경범죄처벌법위반	경범죄처벌법위반

Table 36: List of query case types for IT or Privacy offenses

Law	Crime
개인정보보호법	개인정보보호법위반
통신비밀보호법	통신비밀보호법위반

Table 37: List of query case types for Bribery

Law	Crime
형법 제129조 제1항	뇌물수수
형법 제133조 제1항	뇌물공여
특정범죄가중처벌등에관한법률 제2조, (형법 제129조, 제130조, 제132조)	특정범죄가중처벌등에관한법률위반(뇌물)

Table 38: List of query case types for Intellectual property rights

Law	Crime
상표법	상표법위반
저작권법	저작권법위반
특허법	특허법위반

Table 39: List of query case types for Military duty law

Law	Crime
병역법	병역법위반
군형법	군형법위반

Table 40: List of query case types for Car-related law

Law	Crime
자동차손해배상보장법	자동차손해배상보장법위반
자동차관리법	자동차관리법위반