

論 文 内 容 の 要 旨

博士論文題目

Information Retrieval and Natural Language Inference in the Legal
Domain

(法律分野における情報検索と自然言語推論)

氏 名 CHOI Jungmin

(論文内容の要旨)

Law is one of the domains where the application of natural language processing is expected to bring immense benefit to society. It is an important societal task to lower the barrier to accessing legal services and facilitate the function of law throughout society. We ultimately aim to build a system that assists users untrained in law to find information on their legal issues. To this end, we study two fundamental tasks, information retrieval and natural language inference.

In the first chapter, we survey previous work on natural language processing in the legal domain. Previous work can be divided into the following four broad categories, according to each part of legal professionals' work that is being automated: judgment

prediction, information extraction, question answering, and summarization. Among these, we find information retrieval, particularly law retrieval, and question answering, particularly legal natural language inference, to be the most immediately useful from a layperson's perspective, and therefore set these as the focus of our interest. In the subsequent two chapters, this dissertation presents two essential subtasks, law retrieval and legal natural language inference. Law retrieval is the task of retrieving relevant laws given the query describing a specific situation. The main challenge is recognizing semantic similarity in sentences written in legal and ordinary language. We address this issue by incorporating the hierarchical structure of law in the sentence embedding space by contrastive learning. Legal natural language inference is the task of predicting if a hypothesis on a legal issue is entailed by the premise consisting of laws. Again, the difference in language interferes with the model's ability to recognize entailment. In addition, the existence of multiple rules in the premise poses the unique challenge of identifying the rule that applies to the situation in the hypothesis. We solve this issue by employing a generative model to generate the consequence of the condition. In the final chapter, we make a conclusion and discuss future work.

氏 名	CHOI Jungmin
-----	--------------

(論文審査結果の要旨)

Law is one of the domains where the application of natural language processing is expected to bring immense benefit to society. It is an important societal task to lower the barrier to accessing legal services and facilitate the function of law throughout society. The research in this thesis ultimately aim to build a system that assists users untrained in law to find information on their legal issues. To this end, the thesis study two fundamental tasks, information retrieval and natural language inference.

Law retrieval is the task of retrieving relevant laws given the query describing a specific situation. The main challenge is recognizing semantic similarity in sentences written in legal and ordinary language. This work addresses this issue by incorporating the hierarchical structure of law in the sentence embedding space by contrastive learning. Legal natural language inference is the task of predicting if a hypothesis on a legal issue is entailed by the premise consisting of laws. Again, the difference in language interferes with the model's ability to recognize entailment. In addition, the existence of multiple rules in the premise poses the unique challenge of identifying the rule that applies to the situation in the hypothesis. This work solves this issue by employing a generative model to generate the consequence of the condition.

The research in this thesis starts from careful analysis of the task settings specific to law domains and presents solutions to address their issues, i.e., contrastive learning to represent the hierarchical structure of laws and text generation for complex entailment classification. The proposed method in each task is sound and their effectiveness is demonstrated by carefully designed experiments. The studies are published in one high quality peer-reviewed journal paper and one peer-reviewed international conference paper. The research would have an impact not only to the legal domain of natural

language processing, but to the relevant fields of information retrieval and natural language inference. As a result, the thesis is sufficiently qualified as a Doctoral thesis of Science.