

Japanese Tort-case Dataset

Version: 2024-04-01

Summary

The Japanese Tort-case Dataset (JTD) is built for research on legal judgment prediction in the Japanese language and law. JTD is tailored for tasks, predicting judgments in tort cases and extracting the rationales behind those judgments. The rationale extraction task involves extracting arguments from the claims made by either the plaintiff or the defendant that were important in the court's judgment on the tort case. JTD is constructed from 3,477 civil case judgments annotated by 41 legal experts, consisting of 7,978 instances (with 59,697 claims included in these instances made by plaintiffs and defendants).

This dataset consists of the following two datasets, both provided exclusively for research purposes. As the terms and conditions are different in the two, please carefully review them before use:

- Japanese Tort-case Dataset (Original Documents Only)
- Japanese Tort-case Dataset (Annotations Only)

Only the development and training data are currently available.

This language resource was funded by JST RISTEX JPMJRX19H3 and JST ACT-X JPMJAX20AM. The original judgment documents were provided by [LIC CO.,Ltd.](#)

Citation

- Dataset paper:
 - Hiroaki Yamada, Takenobu Tokunaga, Ryutaro Ohara, Akira Tokutsu, Keisuke Takeshita, Mihoko Sumida. *Japanese Tort-case Dataset for Rationale-supported Legal Judgment Prediction*. ArXiv Preprint, abs/2312.00480.
- Annotation scheme paper:
 - Hiroaki Yamada, Takenobu Tokunaga, Ryutaro Ohara, Keisuke Takeshita, and Mihoko Sumida. 2022. *Annotation Study of Japanese Judgments on Tort for Legal Judgment Prediction with Rationales*. In Proceedings of the Thirteenth Language Resources and Evaluation Conference, pages 779–790, Marseille, France. European Language Resources Association.

Directory Structure

- README_en.md (this file)
- README_en.pdf (pdf version of this file)
- annotations (Japanese Tort-case Dataset (Annotations Only))
 - annotation_guideline (folder containing guidelines for annotators)
 - test.jsonl (test set, to be released later)
 - train.jsonl (training set)
 - valid.jsonl (development set)
- LIC_documents (Japanese Tort-case Dataset (Original Documents Only))
 - doc_id_list (list of judgment document IDs according to the LIC's database)

- test (original documents used in the test set, to be released later)
- train (original documents used in the training set)
- valid (original documents used in the development set)

File Specifications

annotations (Japanese Tort-case Dataset (Annotations Only))

- This file contains formatted annotation data compatible with the huggingface dataset.
- File format: UTF-8, LF, json.
- Specifications of each line in the jsonl file:
 - instnace_ID: ID
 - CD_text: text describing judgment on the tort by the court (not used in the tasks)
 - CD_decision: T_g , judgment on the tort by the court
 - UF_text: U , undisputed facts
 - Date_AD: Date (Gregorian calendar)
 - Category: Case categories (not used in the task)
 - Court: Court in charge
 - FC_text_plaintiff: P , plaintiff's claims (Factual Claims)
 - FC_labels_plaintiff: R_g^P , court's acceptance or rejection of FC_text_plaintiff
 - FC_text_defendant: D , defendant's claims (Factual Claims)
 - FC_labels_defendant: R_g^D , court's acceptance or rejection of FC_text_defendant
 - FC_text_other: Not used
 - FC_labels_other: Not used
 - NC_text_plaintiff: P , plaintiff's claims (Claims of Norms)
 - NC_labels_plaintiff: R_g^P , court's acceptance or rejection of NC_text_plaintiff
 - NC_text_defendant: D , defendant's claims (Claims of Norms)
 - NC_labels_defendant: R_g^D , court's acceptance or rejection of NC_text_defendant
 - NC_text_other: Not used
 - NC_labels_other: Not used

P in the task is provided as a list combining FC_text_plaintiff and NC_text_plaintiff. R_g^P also combines FC_labels_plaintiff and NC_labels_plaintiff. The same applies to D and R_g^D .

LIC_documents (Japanese Tort-case Dataset (Original Documents Only))

- File format: UTF-8, LF, plain text

Additional Resources (for future update)

- <https://github.com/nightowl-jp/JTD-public>