研 究 終 了 報 告 書

「民事紛争のための説明可能な解決結果予測モデル」

研究期間: 2020 年 12 月～2023 年 3 月
研 究 者: 山田 寛章
加速フェーズ期間: 2023 年 4 月～2024 年 3 月

1. 研究のねらい

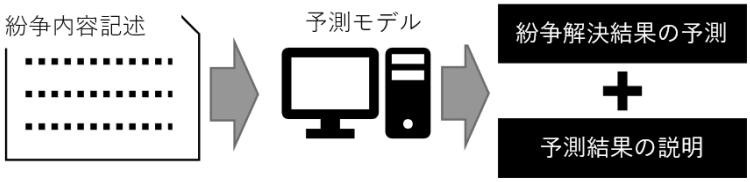

図 1: 説明可能な紛争解決結果予測システムの概要

本研究は、自然言語処理技術を利用し、民事紛争内容の記述から紛争解決の予測結果とその根拠となる説明を与えるモデルの開発を行う。民事紛争において、非専門家の当事者が状況を整理した上で紛争解決の結果を予見することは困難である。紛争内容からその解決結果を予測するモデルが実現できれば、非専門家であっても紛争当事者が状況を自己診断することや、紛争当事者の話し合いにより合意することで紛争の解決を図る調停手続きの進行をモデルによって支援することが可能となる。

実社会において利用するモデルにはその出力の根拠が説明できる必要があり、本研究で対象とするモデルのようにその用途が実際の法的手続きと密接に関わる場合、予測結果に至った根拠や対応する事実関係を同時に提供することで、事後に専門家がその結論を検証可能であることが求められる。

そこで本研究では、民事紛争において事実関係および当事者の主張から紛争解決結果を予測し、その予測結果に至る根拠や説明を併せて出力可能なモデル（図 1）の開発を目的として、課題 A「紛争解決結果予測タスクの定式化とデータセット構築」及び課題 B「説明可能な紛争解決結果予測モデルの開発」に取り組む。

課題 A では、本研究が目指す「解決結果予測」と「説明」の出力とその対となる入力の要素・形式を法学研究者の助言を得て有用性の検討を行い、定式化を行う。自然言語を入力とする設計の紛争解決結果予測タスクは、日本法制度下では先例がない。このため、既存のデータセット・先行するタスク設計が存在せず、入出力形式を設定含むタスク定式等の基礎的検討が不可欠である。その上で、機械学習によるアプローチにも耐えうる数千事例規模の大規模データセット構築を行う。

課題 B では、構築したデータセットを用いて機械学習による紛争解決結果予測タスクと説明タスクの実験を行う。両タスクは、入力の同じ箇所を参照することが予想でき、各タスクのモデルの学習内容を共有できれば、各タスクに対して別個にモデルを構築するよりも性能向上が見込める。本研究では、入力の紛争内容記述を処理するエンコーダを全タスクで共有しながら同時学習する深層学習モデルを構築する。 このエンコーダには BERT に代表される事前学習済みモデルが使用可能であるが、現状法律分野向けに訓練されたものは存在しないため、本研究では法律分野に特化した事前学習済みモデルの開発も行う。

加速フェーズでは、これまでに構築した研究基盤を基により実用シナリオに添った研究展開を行うべく、予測性能の向上(XA)・性能評価方法の改善（XB）・プロトタイプ実装(XC)・データセット利用

促進及び拡張(XD)を主な課題とする。課題 XA「主張間関係を考慮した紛争内容記述エンコーダの改良」:これまでの実験で、双方の主張を考慮可能なエンコーダを利用する手法が有用であったものの、精度にはまだ向上の余地が大きく残っているため、その改善が必要である。また、大規模言語モデルを用いることで各タスクの性能を向上させることを検討する。課題 XB「法律専門家による人手ベースラインスコアの確立」:これまでの人間によるモデル性能分析では、難事例を抽出した上での事後分析に終始していた。現状のモデル群が人間の専門家と比較してどの程度の能力を持つのか明らかにするため、比較対象として人間の専門家がタスクを解いた際のスコアを用意する。課題 XC「紛争解決結果予測モデルを用いたプロトタイプ」:データセットにはない未知の入力を試したいというニーズ、研究者ではない第三者へのデモンストレーションを想定すると、実働するアプリの構築が必要となる。本研究で構築したモデル及びデータセットを元にデモアプリを作成する。課題 XD「JTD の利用・応用促進」:成果 A-2 のデータセット JTD は当初計画を大きく上回る規模・質となっており、今後の法情報学領域における基盤的データセットとなり得る。例えば、JTD を活用したコンペティション/ワークショップの開催により利用促進を図ると共に、本研究分野の振興を目指す。

2. 研究成果

(1)概要

本研究の中心的研究課題として、課題 A「紛争解決結果予測タスクの定式化とデータセット構築」及び課題 B「説明可能な紛争解決結果予測モデルの開発」が存在する。

課題 A では、法学研究者との意見交換を通じて次のようにタスクを設定した。本研究計画で対象とする判決書を不法行為について判断している事件に限定する。モデルへの入力は、ある不法行為についての訴えに関する前提事実・原告側の主張・被告側の主張とし、出力はその不法行為が裁判所によって認められたか否かとする。モデルがある不法行為が成立するか否かについて予測したことの根拠として、入力された原告・被告の主張群の中から判断の根拠となった主張を抽出することで説明性を担保する。

さらに、日本法下として初めての専門家によるアノテーション付き判決書大規模データセット(約 8000 事例収録)の構築を行った。言語・法制度にかかわらず、この規模で専門家によるアノテーションを実施した類例はない。さらに、判決結果アノテーションに留まらず、判決結果説明として裁判当事者(原告・被告)の主張が裁判所の判断中で採用されたか否かがアノテーションされている。結果として、国外先行研究のデータセットと比べてアノテーション信頼性・説明タスク向けアノテーションの有用性の点で優位であり、今後の紛争解決結果タスク向けデータセットのデファクトスタンダードを狙えるものとなっている。

課題 B では、1)判決書に特化した事前学習済言語モデルの構築、2)根拠抽出と解決結果予測実験を行った。

近年の自然言語処理研究においては、大量の文書で半教師あり学習を行う事前学習済言語モデルをベースラインモデルとして用いたり、発展的なモデルの開発基盤として用いたりすることが一般的であるが、日本語かつ法律分野に特化した事前学習済モデルがこれまで存在しなかった。そこで、日本の民事事件判決書(約 5.4GB)によって事前学習された JLBERT を構築し、その性能評価と分析を行った。

さらに、本研究で示した紛争解決結果予測タスクの Proof-of-concept として、構築したデータセットを用いた根拠抽出と解決結果予測実験を行った。

加速フェーズ期間においては、上記マルチタスク学習アプローチ(提案手法)をより改善することで、解決結果予測タスクにおいて約 68%、根拠抽出タスクにおいては約 65%の Accuracy を達成した。他のベースライン手法との比較の結果、提案手法において根拠抽出と解決結果予測のマルチタスク学習を行うことの有用性を示した。また、法律専門家による上記提案手法の出力結果の一部の人手による分析を行った。提案手法で予測を誤った事例を弁護士・法学研究者によって、個別分析した結果、75%の事例が専門家にとってもその予測に確信が持てない難事例であったことが示された。さらに、本研究で構築したデータセットの各タスクについて、現状のモデル群が人間の専門家と比較してどの程度の能力を持つのか明らかにするため、比較対象として人間の専門家がタスクを解いた際の評価も実施した。

加速フェーズにおいて、本研究での成果の応用・利用を促進するため、構築したデータセット及び提案手法を応用した法律相談デモシステムを構築した。このシステムはデータセットが対象としている発信者情報開示請求及びそれに関連する不法行為事件を中心とした相談にしか対応していないものの、大規模言語モデルを組み合わせることで自然言語による流ちょうなユーザ体験が可能となっている。

(2)詳細

1. 紛争解決結果予測タスクの定義 (成果 A-1)

法学研究者との意見交換の結果、本研究での実験対象を不法行為について判断している事件に限定した。不法行為は民事事件で頻出の概念であり、権利侵害とその損害賠償請求に関わる。近年件数が増加しつつあるインターネット上での権利侵害(例:名誉毀損)も不法行為の一類型であり、紛争解決や調停支援の自動化対象として法学的観点から意義深い。不法行為は民法の中でも「一般条項」という形で定義される。具体的な要件が定義されず、抽象的な価値基準として示されているため、従来の紛争解決結果予測で用いられてきた論理プログラミングやルールベースによる手法では解決が難しい。このため、本研究課題のとる機械学習アプローチによって初めて解決できる課題であり、技術的にも挑戦する意義が大きい。

本研究で開発するモデルは、紛争内容を自然言語で記述した文書を入力とし、「解決結果予測」と「説明」を出力することを想定している。当初の研究計画では、タスク候補の一つとして紛争解決結果・説明それぞれにおいて、判決文生成・説明文生成を検討していた。しかし、法学専門家を交えたタスク設計検討の結果、不法行為の予測であ

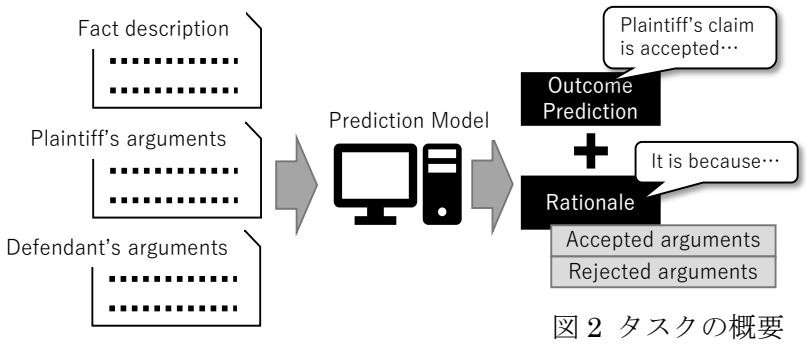

図 2 タスクの概要

れば二値分類で自明であること、判決書における説明は裁判所によってどの主張が採用され・棄却されたのかに依る部分が大きいため新たなテキストを生成する必要性が薄い、との結論を得たため本研究の範囲では文生成は見送った。各タスクの形式は次のような形式となった。モデルへの入力はある不法行為についての訴えに関する前提事実（争いのない事実）・原告側の主張・被告側の主張とし、出力はその不法行為が裁判所によって認められたか否かとする。不法行為の成否をモデルが予測したことの根拠として、入力された原告・被告の主張群の中から判断の根拠となった主張を抽出することで説明性を担保する。（図 2）

2. 判決書大規模データセットの構築(成果 A-2)

このようなタスクを解くモデルの訓練・評価ためのデータセットを構築した。データセット構築に際し、判決書中から不法行為に関する主張や事実のみを抽出する必要と、不法行為が成否の記述（≒答え）を含んでいる「裁判所の判断」節のテキストがモデルへの入力としてデータセットに収録されて教師信号のリークが起こらないように配慮する必要があったため、法律知識を持った作業者による人手アノテーションが必須であった。また、本研究計画では機械学習による予測モデルの構築を目指していることから、可能な限り多くの事例数を収集する必要があった。これらの要件から、多数の専門家による同時進行での大規模アノテーションを実施する野心的なデータセット構築を行った。

多数の作業者が同時進行でアノテーションを行っても、作業者間で一貫した基準のアノテーションが実施できることを担保するため、ガイドラインやチュートリアル等の整備を行った。この際、5 人程度の小規模な作業者グループにて 5～25 件程度の少ない文書を用いて予備的なアノテーションを３回実施した。予備アノテーションでは、各回で Fleiss’s κ や Krippendorff’s αU などの指標を用いて作業者間一致度を測定し、ガイドラインの改定・チュートリアルの整備やサンプルアノテーションの拡充によって作業者間での作業内容の一貫性が改善されていったことを確認した。3 回目の予備アノテーションの結果として αU=0.654 の値を記録しており、開発したアノテーションスキームで安定した作業を行えることが確認できた。[成果リスト P1]

アノテーション作業は、ある不法行為に関連する「争いのない事実」の抽出、「原告の主張」「被告の主張」の抽出及び各主張が裁判所によって認められたか否か、そして、裁判所の判断としてその不法行為が成立したのか否かの判定が主な内容となる。また、一件の判決書中で異なる複数の不法行為について判断されている場合があるため、これらを区別して特定の不法行為と特定の主張や事実との関係を対応づけるためのアノテーションも行った。このアノテーションの結果として図 3 のようなデータが作成される。最終的には図 3 のよう

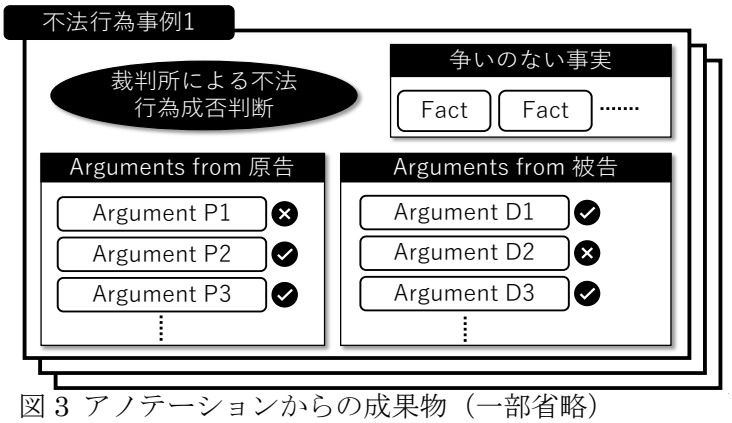

図 3 アノテーションからの成果物（一部省略）

な事例データが不法行為事例の数だけ得られる。

日本法下として初めての専門家 47 名（法曹等専門家：11 名、法科大学院卒業生：25 名、法科大学院生等：11 名）によるアノテーション付き 8132 事例収録の日本語不法行為データセット(Japanese Tort-case Dataset; JTD)の構築を行った。JTD は、当初の研究計画での想定成果を質（計画では主に法学部生によるアノテーション＋自動ラベル付与）・量（計画時は約 1000 事例）共に大きく上回った。言語・法制度にかかわらず、この規模で専門家によるアノテーションを実施した類例はない。さらに、判決結果アノテーションに加えて、判決結果説明として裁判当事者（原告・被告）の主張が裁判所の判断中で採用されたか否かがアノテーションされている。国外先行研究のデータセットと比べても、専門家によるアノテーションの信頼性・説明タスク向けアノテーションの有用性の点で優位であり、今後の紛争解決結果タスク向けデータセットのスタンダードを狙えるものとなっている。[成果リスト R1]

3. 判決書に特化した事前学習済言語モデルの構築（成果 B-1）

近年の自然言語処理研究では、大量の文書を用いて半教師あり学習を行う BERT のような大規模な事前学習済言語モデルを、ベースラインモデルとして用いたり発展的なモデルを開発するための基盤として用いたりすることが一般的になっている。英語においては、オリジナルの BERT のように Wikipedia を用いて訓練された事前学習済モデルの他に、法律分野に特化して訓練された BERT も存在しており、法律関連の自然言語処理タスクにおいてオリジナル BERT よりも高い性能を達成していることが報告されている。一方、日本語においては法律分野に特化した事前学習済モデルがこれまで存在しなかった。そこで、国内で初めてとなる、日本の民事事件判決書 170,320 件（約 5.4GB)によって事前学習された JLBERT-SC 及び、Wikipedia と判決書の両方を用いた事前学習した JLBERT-FP を構築した。構築した JLBERT の性能を判決書重要箇所抽出タスクによって評価したところ、JLBERT-FP/SC 共に通常の日本語 BERT と比較して性能向上が確認された。[成果リスト P2, P3, P4]

4. 根拠抽出と解決結果予測実験（成果 B-2）

さらに、本研究で設計した各タスクの概念実証として、構築した JTD を用いて根拠抽出と解決結果予測実験を行った。

5. 加速フェーズにおける成果

マルチタスク学習アプローチ（提案手法）をより改善することで、解決結果予測タスクにおいて約 68%、根拠抽出タスクにおいては約 65%の Accuracy を達成した。他のベースライン手法との比較の結果、提案手法において根拠抽出と解決結果予測のマルチタスク学習を行うことの有用性を示した。また、法律専門家による上記提案手法の出力結果の一部の人手による分析を行った。提案手法で予測を誤った事例を弁護士・法学研究者によって、個別分析した結果、75%の事例が専門家にとってもその予測に確信が持てない難事例であったことが示された。さらに分析を進めたところ、難事例のうち 86.7%について事例を解くには追加の情報が必要だったということが分かった。この 86.7%の内訳は、本来必要な事例固有の情報が入力としての情報の中から欠落しているケース（情報不足, 58.3%）、法律分野特有の知識が

必要なケース（法的分野知識, 25.0%）、法律分野特有ではない一般常識的な認識が必要なケース（一般常識, 3.3%）、及びその他であった。この分析結果は、構築した現状のデータセットの限界を明確にし、さらに将来のアノテーション改善の指針の策定に有用な情報をもたらした。［課題 XA］

さらに、本研究で構築したデータセットの各タスクについて、現状のモデル群が人間の専門家と比較してどの程度の能力を持つのか明らかにするため、比較対象として人間の専門家がタスクを解いた際のスコアを準備している。作業対象は約 200 件を予定している。本報告書追記時点（12/1）で専門家による作業途中であり、今後そのとりまとめと分析を行う。［課題 XB］

加速フェーズにおいて、本研究での成果の応用・利用を促進するため、構築したデータセット及び提案手法を応用した法律相談システムのデモを構築した。このシステムはデータセットが対象としている発信者情報開示請求及びそれに関連する不法行為事件を中心とした相談にしか対応していないものの、大規模言語モデルを組み合わせることで自然言語による流ちょうなユーザ体験が可能となっている。［課題 XC］

本研究において構築したデータセット JTD は当初計画を大きく上回る規模・質となっており、今後の法情報学領域における基盤的データセットとなり得る。例えば、JTD を活用したコンペティション/ワークショップの開催により利用促進を図ると共に、本研究分野の振興を目指す。ACT-X の研究成果の総括として、構築したタスク・データセット・ベースラインモデルを利用したリーダボードシステムの構築を行っており、年度末までに作業が完了する見込みである。ACT-X の研究期間終了後も、このリーダボードシステムを活用して、研究成果活用の最大化を目指す。［課題 XD］

6. ACT-X 研究領域内外・産業界との連携

法制度への人工知能実装の可能性を探るプロジェクトである「法制度と人工知能（JST RISTEX/HITE, 研究代表者：角田美穂子, グラント番号：JPMJRX19H3）」の研究グループとの研究協力を行い、アノテーション作業者の募集、アノテーションスキームの開発、本研究にて開発した予測モデルの出力の評価において相互協力した。また、本研究のアウトリーチ活動の一環として、一橋大学法学部の集中講義「テクノロジーとリーガルイノベーション」にゲスト講師として参加し、本研究活動の概要を紹介した。これらの ACT-X 領域外連携の成果として、法学における AI・その他情報技術応用の前提としてのデータの重要性を指摘した他、データセット構築手法・データ共有のための規約制定のノウハウといった、法-情報間の分野横断的な研究を促進するための基盤を構築してきた。［成果リスト P5, P6］

産業界との連携として、本研究での民事裁判例データ提供元である株式会社 LIC とは、判決書からの重要箇所抽出及び自動要約システムの研究開発について 2022 年 9 月から共同研究を実施することとなった。本研究における法律特化 BERT 構築手法及び、データセット構築に関して得られた知見からの発展的成果となる。

3. 今後の展開

政府のオンライン裁判外紛争解決手続き（ODR)活性化検討会による、ODR 活性化とりまとめ報

告[1]では、身近に法律家がおらず司法アクセスが容易でない「司法過疎」問題が潜在することが指摘されている。今後の超高齢社会の進展によっては司法過疎が一層深刻化することが予想される。本研究での成果を応用し、当事者に法的紛争の予測結果を提供することで、この問題を解消することが可能となる。紛争当事者の話し合いにより合意することで紛争の解決を図る調停手続きを予測モデルによって支援することや、モデルを ODR に組み込み定型的な紛争調停を自動化・省力化することによって、いつでも/どこでも/だれでも容易に司法へアクセス可能となることが、最終的な目標である。

そのような目標を実現するためには、①データ等研究資源の拡充、②予測精度の向上、③社会に受容される実装、を達成しなければならない。

[1]https://www.kantei.go.jp/jp/singi/keizaisaisei/odrkasseika/pdf/report.pdf

4. 自己評価

① 研究目的の達成状況

課題 A「紛争解決結果予測タスクの定式化とデータセット構築」では、計画を大きく上回る規模と質のデータセットを構築した。課題 A の目的であったタスク定式化とデータセット構築は十二分に達成できたと評価できる。

課題 B「説明可能な紛争解決結果予測モデルの開発」では、機械学習による紛争解決結果予測タスクと説明タスクの実験実施による概念実証と、法律分野に特化した事前学習済みモデルの構築が目的であった。報告書執筆時点では、紛争解決結果予測タスクと説明タスクのマルチタスク学習によるモデルを含む複数の手法を用いた実験が進行中であり、一部の結果が出そろっていない。しかし、研究期間終了までには計画を完遂できる見込みである。法律分野に特化した事前学習済みモデルの構築では、モデルの性能分析について興味深い結果を得られたため、独立して論文化した。

総じて、A・B 両課題について当初の研究目的は達成できる見込みであり、一部については当初計画を上回る成果を得られた、と評価できる。

（加速フェーズ実施後追記）

加速フェーズで改良した予測モデルの実験結果とその結果の専門家による人手分析の内容は、従来の研究成果とまとめてジャーナル論文[P8]として採択されており、研究プロジェクト全体の総括的な成果となった。また、加速フェーズで新規に実施した専門家によるベースラインスコアの取得・デモ構築・JTD の利用促進についても、完遂の目処がたっており、加速フェーズ期間中の課題についても当初の目的を達成できたと評価できる。

② 研究実施体制及び研究費執行状況

実施期間中、研究者の所属が学生、ポスドク、助教と毎年変更となったものの、エフォートや研究実施体制への影響はほとんどなく、期間を通じて研究実施体制に問題は発生しなかった。特にポスドク以降は、研究者としての独立性が高まり、予算執行や研究実施がよりスムーズに進んだ。研究費執行については、初年度（12 月開始のため）を除き、年度後半に執行が集中しないよう計画的な執行を実施できた。

③ 研究成果の科学技術及び社会・経済への波及効果

本研究の成果物である日本語不法行為データセット(JTD)は、約 8000 事例を収録している。

この規模で法の専門家によるアノテーションを実施した研究は、国際的にも例がない。欧州・米国・中国における先行研究の多くは、裁判所等から提供されたメタデータを利用した判決（勝訴 vs 敗訴）レベルの粒度の粗い自動アノテーションであったり、根拠等を別途アノテーションしていても単独の作業者のみによるものだったりと、ラベルの信頼性に疑問が残る。JTD では、複数の専門家によるアノテーションを作業者間一致度の検証を行った上で実施しており、より信頼性が高いといえる。本研究で開発した構築手法が、国際的にも今後の法情報学におけるデータセット構築の例として参照されることが期待できる。

本研究で開発した予測モデルは、非法律専門家の紛争解決結果の予見可能性向上と紛争調停手続きの効率化を提供するという点で社会的に意義深い。政府のオンライン裁判外紛争解決手続き（ODR)活性化検討会による「ODR 活性化とりまとめ報告」にも、身近に法律家がおらず司法アクセスが容易でない「司法過疎」問題が潜在することが指摘され、今後の超高齢社会の進展によって司法過疎問題が一層深刻化していくことが懸念されている。紛争当事者の話し合いと合意によって紛争解決を図る調停手続きのモデルによる支援や、モデルを ODR に組み込むことによる定型的な紛争調停の自動化・省力化によって、いつでも/どこでも/だれでもが容易に司法へアクセス可能となる社会の実現を支える。

5. 主な研究成果リスト

（1）代表的な論文（原著論文）発表

研究期間累積件数：2件

研究期間累積件数：3件（加速フェーズ実施後更新）

1. Hiroaki Yamada, Takenobu Tokunaga, Ryutaro Ohara, Keisuke Takeshita, Mihoko Sumida. Annotation Study of Japanese Judgments on Tort for Legal Judgment Prediction with Rationales. Proceedings of the Thirteenth Language Resources and Evaluation Conference. 2022, 1, pp. 779-790 [P1]
日本法下における法的判断予測モデル研究に向けたアノテーション済コーパスを構築すること目指し、独自のアノテーションスキームを提案した。このスキームは裁判所の判断とその根拠を抽出でき、同時に各裁判所の判断と対応する根拠の間の因果関係を紐付けることが可能な設計である。不法行為事件を題材としたアノテーション作業者間一致度測定結果によれば、提案スキームにより品質を確保したアノテーションを実施できることが確認された。
2. Keisuke Miyazaki, Hiroaki Yamada and Takenobu Tokunaga, Cross-domain Analysis on Japanese Legal Pretrained Language Models. In Findings of The 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing. 2022, in press. [P2]
本論文では日本法分野に特化した事前学習済言語モデル(PLM)の構築方法の提案と、その性能分析結果を報告する。①一般ドメインで事前学習後に法ドメインで追加事前学習することで PLM の性能が向上する、②法特化 PLM は法ドメイン特有の語義と一般ドメインにおける語義の両方の認識を両立し、区別できる、③追加事前学習で構築された法特化 PLM は法ドメインでの性能が向上するが、法以外のドメインにおける性能も確保される。
3. Hiroaki Yamada, Takenobu Tokunaga, Ryutaro Ohara, Akira Tokutsu, Keisuke Takeshita,

Mihoko Sumida, Japanese Tort-case Dataset for Rationale-supported Legal Judgment Prediction. ArXiv, preprint. 2023. [P7] / Hiroaki Yamada, Takenobu Tokunaga, Ryutaro Ohara, Akira Tokutsu, Keisuke Takeshita, Mihoko Sumida, Japanese Tort-case Dataset for Rationale-supported Legal Judgment Prediction. Artificial Intelligence and Law, to appear. 2024. [P8]

本論文では、判決予測(LJP)のための初のデータセットである日本不法行為判例データセット(JTD)を提案する。JTD には、不法行為予測、根拠抽出の 2 つのタスクがある。根拠抽出タスクは、原告と被告が提出した主張の中から裁判所が認めた主張を特定するもので、この分野における新しいタスクである。JTD は 41 人の法律専門家によってアノテーションされた 3,477 件の民事裁判例から構築されており、7,978 の事例とそれらに内包される 59,697 の主張を含んでいます。ベースライン実験では、提案した 2 つのタスクの feasibility が示されたた。また、法律専門家による分析を通してエラーの原因解明も行った。

（2）特許出願

該当なし

（3）その他の成果（主要な学会発表、受賞、著作物、プレスリリース等）

国内会議プロシーディングス（査読なし）

1. 宮崎桂輔, 菅原祐太, 山田寛章, 徳永健伸. 日本語法律分野文書に特化した BERT の構築. 言語処理学会第 28 回年次大会発表論文集. 2022, 1, pp. 1546-1551 [P3]
2. 菅原祐太, 宮崎桂輔, 山田寛章, 徳永健伸. 日本語法律 BERT を用いた判決書からの重要箇所抽出. 言語処理学会第 28 回年次大会発表論文集, 2022, 1, pp. 838-841 [P4]
3. 伊藤光一, 山田寛章, 徳永健伸. 低資源な法ドメイン含意タスクにおけるデータ拡張. 言語処理学会第 29 回年次大会発表論文集, 2023, 1, pp. 990-995
4. 山田寛章, 徳永健伸, 小原隆太郎, 得津晶, 竹下啓介, 角田美穂子. 日本語不法行為事件データセットの構築. 言語処理学会第 30 回年次大会発表論文集. 2024, 1, pp. 1045-1050.（委員特別賞）
5. 山田寛章, 小原隆太郎, 角田美穂子. 不法行為判断予測データセット構築における ELSI 課題 人工知能学会全国大会論文集. 2024. To appear.

国内学術雑誌（査読なし）

1. 山田寛章. 法と人工知能の接点. 情報法制研究. 2022, 11 巻, pp. 27-33 [P5]

口頭発表

1. 山田寛章. 日本語判決書を用いたデータセットの構築. 言語処理学会第 28 回年次大会併設ワークショップ JED2022 日本語における評価用データセットの構築と利用性の向上. 2022. [P6]
2. 山田寛章, 角田美穂子. Japanese Tort-case Dataset for Rationale-supported Legal Judgment Prediction. International Research Project “Legal Systems and Artificial Intelligence” Final Events. 2023

データセット

1. Hiroaki Yamada, Mihoko Sumida. Japanese Tort-case Dataset. [R1]