

法令沿革の自動作成における予測結果の誤り検出手法の提案

前原太陽¹ 佐野智也² 竹中要一³

¹ 関西大学 ² 名古屋大学 デジタル人文社会科学推進センター

³ 関西大学 総合情報学部

¹suisougakuperc@gmail.com ²tomoya@law.nagoya-u.ac.jp

³takenaka@kansai-u.ac.jp

概要

我々は条文の内容理解を促進するための条文沿革を自動で生成する手法を研究してきた。しかしながら、予測する結果が手作業で対応付けした資料上の正解と完全に一致させることは困難であるため、法学研究のような厳密性を求める場合には、コンピュータの予測結果の整合性を確認する作業が必要な状況である。本研究では、コンピュータの予測結果の整合性を確認する作業工数を軽減するための手法について提案する。旧条文において条番号が連番になっている条文同士は新条文においても連番になっているという傾向を利用し、コンピュータで予測した条番号のような連続値を、誤り検出のための入力値として利用する手法を提案する。実験には昭和13年に改正された商法を用い、ドットマトリクスを利用して考察を行う。

1 はじめに

法学研究および法律実務において、条文の解釈を行う際には、当該条文がどのような立法趣旨や制度的背景の下で制定・改正されてきたのかを把握することが重要である。しかしながら、条文沿革を把握するツールは十分整備されていない。近年、法令情報のデジタル化が進み、条文検索や改正履歴の確認は以前と比べて容易になりつつある。例えば、政府が提供する法令検索サービス「e-Gov 法令検索」¹⁾では、施行中の法令や近年の改正履歴を確認することができる。しかしながら、確認できる主な法令は、サービス開始の2017年以降を対象としており、過去の法律については十分に網羅されていないという課題がある。また、過去の法令情報を整備する取り組みとして、佐野による条文沿革データベース²⁾も

存在するが、現時点では一部の法律に限定された整備にとどまっている。

条文沿革の作成作業について、作業効率化を支援するためにはコンピュータでの自動化が必須である。現在、法学研究者や法律実務者が条文沿革を把握しようとする場合、複数の資料を参照しながら手作業で対応関係を整理する必要があり、大きな負担が生じている。日本の法律は「e-Gov 法令検索」に掲載されている現行法令だけでも8,931法令あり[1]、改正も継続的に行われているため、条文沿革の作成や更新を手作業のみで行うことは、工数の面で限界がある。したがって、条文沿革作成作業の効率化を図るためには、コンピュータによる支援が不可欠である。

本研究では、条文沿革の自動的な紐づけおよび編集補助機能を備えたツールの開発を最終的な目的とする。我々は、自然言語処理技術を用いて新旧条文間の対応関係を自動的に推定する研究に取り組んできた[2, 3]。条文の類似度に基づいて新旧の対応関係を推定する手法を提案し、専門家が手作業で対応付けした新旧条文の対応関係について、90.8%の精度で自動的に条文の紐づけが可能であることを示した。しかしながら、推定結果には、国会で議論される際に利用する新旧対照表とは異なる対応関係を出力するものもあり、法学研究者がそのまま利用することは難しい。現時点の提案モデルでは、人間が全推定結果を確認し、妥当性を判断する必要があるため、確認作業の負担は依然として残されている。

本研究では法学研究者が条文沿革を容易に把握できるようなツールの開発に向け、今までに提案してきた新旧条文の対応推定の結果を基に、推定結果の誤りを検知する手法を提案する。

1) <https://www.e-gov.go.jp/elaws/>

2) <https://law-platform.jp/>

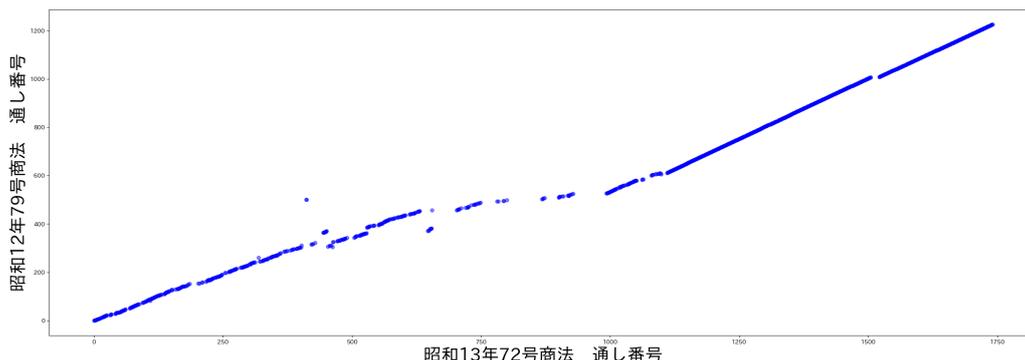


図 1: 昭和 12 年 79 号から昭和 13 年 72 号における新旧条文の対応関係

2 提案手法

本報告の提案範囲は図 2 の赤枠部分であり、コンピュータによって予測された新旧条文の対応関係が正しいかどうかを判定する手法を提案する。モデル全体の入力および出力は下記の通りである。

Input:

- 旧法令のテキスト全行（テキストの集合）
- 新法令の 1 行（テキスト）

Output:

- 入力の新法令の 1 文に対応する旧条文の 1 文の番号
- 対応が正しいか否かの判定結果（bool 値）

我々が今までに提案したコンピュータでの予測は、条文の分散表現を用いた類似度計算に基づき、対応条文を予測してきた。本報告では条番号の連続性という構造的な特徴に着目する。法令の改正においては、条文の内容が引き継がれる場合、新旧条文の番号が連続的に対応することが多い傾向がある。例えば、昭和 12 年 79 号から昭和 13 年 72 号への変遷では、旧法の第 281 条の後継が新法の第 518 条、旧法の第 282 条の後継が新法の第 519 条といった形で、対応関係がほぼ単調に推移するケースが観察される（図 1）。したがって、予測結果の中で特定の条文のみが周囲の条文の番号と大きく異なる対応関係を示す場合、その予測は誤っている可能性が高いと考えられる。対応関係の推移に関する仮説をもとに、コンピュータによる予測結果を用いて、予測結果の整合性検証を行う。

本研究では対象条文の予測結果と、その前後の条文における予測結果との差に着目する。前後条文と

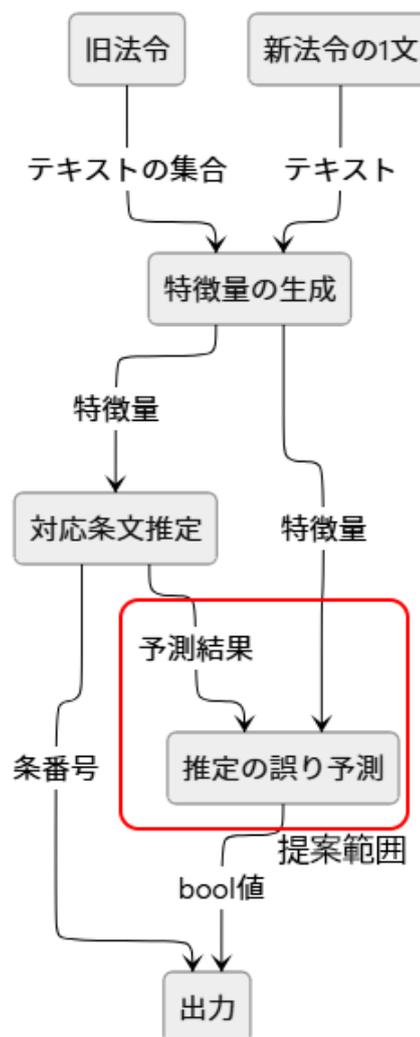


図 2: 処理フロー全体

の予測結果の差を特徴量として用いることで、対象条文の対応関係が外れ値となっていないかを検出できるのではないかと考えた。本手法は、対応関係そのものを再推定するのではなく、既存の推定結果に対して妥当性を評価する点に特徴がある。

3 実験条件

コンピュータで対応条文を予測する方法としては、前原らの手法 [3] における第 2 フェーズを利用する。

過去の実験では、新旧対応の紐づけを条文単位行ってきた。しかしながら、法令が改正されるタイミングで異なる条文の一部がまとまって 1 つの条文に集約されることや、分解されることが発生する。したがって、今回の実験では、従来の条文単位での分析ではなく、法令に記載されている「行」ごとに行う。本来、法令上は「行」ごとの通し番号がないため、前処理として、1 行ごとに通し番号を採番する。なお、行単位の紐づけにおいては対応関係に関する答えが存在しないが、対応する通し番号が属する条番号が正解の条番号と一致していれば、正解と考える。

文章の分散表現の生成には Sentence-BERT[4] を用いる。学習済みモデルとしては、日本語文書に対して高い性能が報告されている `sonoisa/sentence-bert-base-ja-mean-tokens-v2`³⁾ を利用する。

実験対象として、昭和 12 年法律第 79 号から昭和 13 年法律第 72 号への商法改正を取り上げる。昭和 13 年法律第 72 号による改正では、条番号の振り直しが発生しており、新旧条文の対応関係を整理する作業が特に多く必要とされた時期である。

提案手法で用いる特徴量および判定手法は以下の通りである。なお、特徴量の選定にあたっては予備実験を行ったが、使用する特徴量の組み合わせによる精度差は大きくなかった。

- 手法
 - 対応条文予測時：LDA (Linear Discriminant Analysis)
 - 提案手法適用時：決定木 (深さ 3)
- 特徴量
 - 予測対象行の文字数
 - N 番目 ($N=\{1,2,5\}$) に類似度が高い行との類似度

3) <https://huggingface.co/sonoisa/sentence-bert-base-ja-mean-tokens-v2>

- N 番目 ($N=\{1,2,4,5\}$) に類似度が高い行の文字数
- 対象行の文字数を 1 番目に類似度が高い行の文字数で割った値
- 対象行の予測通し番号と、対象行の 1 つ前の予測通し番号との差 (提案手法適用時のみ)

4 結果

コンピュータの予測におけるにおける対応通し番号予測の結果、および今回の提案手法によるコンピュータの予測誤りを推定した結果の混合行列を表 1 および表 2 に示す。

表 1: コンピュータの対応条文予測結果

	Predicted Negative	Predicted Positive
Actually Negative	600	79
Actually Positive	94	969

表 2: 混合行列

	Predicted Negative	Predicted Positive
Actually Negative	8	165
Actually Positive	3	1566

また、コンピュータの予測結果をドットマトリクスで可視化したものに、本提案手法にて誤りを検知できたものを「★」として協調したものを図 3 に示す。コンピュータの予測結果においては、新旧の対応関係が外れ値となっている部分がほとんど False Positive となっており、誤った対応関係を予測していることがわかり、今回の提案手法では誤りのうち 8 件を検出できていることがわかる。

5 考察

本節では、提案手法によって True Negative、すなわちコンピュータによる誤った予測を正しく検知できた事例に着目して考察を行う。

図 3 からは条文の対応関係が連続に推移していない場所の付近でコンピュータの予測が失敗しており、提案手法が予測誤りを検知できていることが確認できる。

実際に検知された条文の組を確認すると、表 3 に示す通り、新旧条文の文言は人間の目で見ても完全に一致もしくは数値の違いがある程度であるとわ

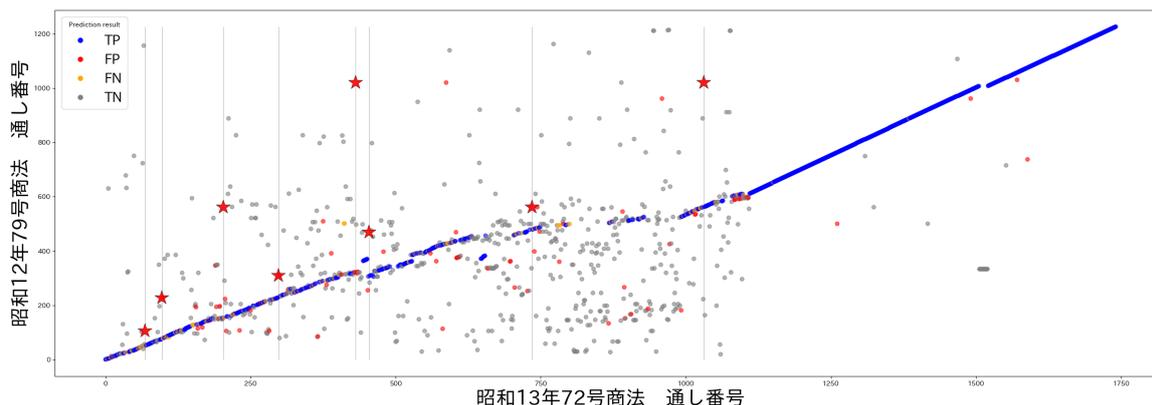


図 3: コンピュータの予測結果と予測誤りの検出に関するドットマトリクス

新法行番号	旧法行番号 (予測)	新条文の文章	旧条文の文章
44.2.1	61-2.2.1	第四十二条第二項の規定は前項の場合に之を準用す	第三十条の二第二項の規定は前項の場合に之を準用す
63.1.2	120.1.2	目的	目的
112.2.1	247.2.1	第八十八条及第九十条第二項の規定は前項の場合に之を準用す	第一百八条第二項の規定は前項の場合に之を準用す
166.1.4	148.1.4	資本の総額	資本の総額
215.3.1	597.2.1	前条第二項の規定は前項の場合に之を準用す	前条第二項の規定は前項の場合に之を準用す
225.1.2	212-3.1.2	会社の商号	会社の商号
353.2.1	247.2.1	第一百八十一条第二項の規定は前項の場合に之を準用す	第一百八条第二項の規定は前項の場合に之を準用す
471.2.1	597.2.1	前条第二項の規定は前項の場合に之を準用す	前条第二項の規定は前項の場合に之を準用す

※行番号が「x.y.z」の場合、「x 条 y 項 z 行目」を指す

表 3: 今回検出できた誤り

かる。表 3 の条文は、法令中で繰り返し用いられる定型的な表現を含んでおり、類似度のみに基づく手法では誤って対応付けられてしまった可能性がある。提案手法は、今回のようなケースにおいて、コンピュータが誤りやすい特徴を適切に検知できていると考えられる。

6 さいごに

本研究では、コンピュータが予測した新旧条文の対応関係に対して、誤りを検知する方法を提案した。対象条文の前後に位置する条文の予測結果を利用することで、対応関係の正誤を 94.9% の精度で推定できることを示した。本手法は、対応関係そのものを自動的に決定するのではなく、人間による確認作業を支援することを目的としている。今後は、これまでに得られた研究成果を GUI として提供し、条文沿革作成作業を支援するツールの開発を進めていきたい。

謝辞

本研究は JSPS 科研費 23K01052 および、2025 年度関西大学学術研究員研究費の助成を受けたものです。

参考文献

- [1] Digital Agency, Government of Japan. Db 登録法令数 — 登録法令情報. <https://laws.e-gov.go.jp/registdb/>, 2025. 参照日: 2025-12-27.
- [2] 前原太陽, 竹中要一, 佐野智也. 国会が告示する改正民法における新旧対応の整合性の検証. 言語処理学会第 31 回年次大会 発表論文集, 2025.
- [3] Taiyo Maehara, Tomoya Sano, and Yoichi Takenaka. Automating the creation of legislative article histories in japanese commercial law: A method for identifying corresponding articles before and after amendments. **New Frontiers in Artificial Intelligence**, pp. 114–129, 2025.
- [4] Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. **arXiv preprint arXiv:1908.10084**, 2019.