

論文

サイバーコート実現に向けた技術検討
—— Neural Network を用いた伏字位置予測 ——A Study for Implementation of Cyber Court;
Predicting Unprintable Words in Proceedings by Neural Network伊藤 篤¹・桐生 雄也²・金澤 政和²・笠原 毅彦³¹ 宇都宮大学大学院工学研究科² 宇都宮大学大学院工学研究科情報システム科学専攻³ 桐蔭横浜大学大学院法学研究科

(2018年3月17日 受理)

I. はじめに

司法システムに ICT (Information and Communication Technology) を導入したものを特に「サイバーコート」と呼ぶ。サイバーコートの導入が必要とされる理由としては、以下の2点が挙げられる。

日本国憲法第32条が保証する「裁判を受ける権利」から、「日本の司法へのアクセス (access to justice)」について伊藤正巳は「憲法上いかに裁判を受ける権利が全ての人間に公平に開かれており、形式的に正義へのアクセスの道が認められていても、事実上それを実行せしめることを阻害する要因は少なくない。」とまとめている。しかし、現実的には、裁判所への距離、交通手段の有無、裁判記録の閲覧など、司法へのアクセスは、地理的・肉体的・金銭的に平等ではない。したがって、地理的、肉体的に裁判所まで赴くことが困難な人や法的知識の少ない人についても

平等に法へのアクセスができなくてはならず、これを解決する手法としてサイバーコートが注目されている。

また、近年、国際化の進展により多国間貿易が拡大し、企業の多国籍化が進んでおり、電子商取引の規模は世界的に拡大傾向にある¹⁾。国際的な企業が増えることで、様々な国々の間での紛争も増加傾向にある。従って、多くの国々が司法システムに ICT を導入し、誰もが司法機関に容易にアクセスできるようにすることで、経済成長を支援する柔軟性とオープン性を実現することを目指している。

サイバーコートの研究は米国の Courtroom21 から始まった。これは1993年にウィリアム・アンド・メアリー大学 (William & Mary Law School) と米国の州裁判所の共同プロジェクトとして開始された²⁾。ここでは、テレビ会議システムを利用してサイバーコートを実現し ICT の効果を判断するために多くの試みが行われた。現在、彼らはこの技術を用いてロースクールでの教育を支援

¹ ITO Atsushi: Professor, Graduate School of Engineering, Utsunomiya University, 7-1-2 Yoto, Utsunomiya, Tochigi 321-8585, Japan

² KIRYU Yuya and KANAZAWA Masakazu: Department of Information System Science, Graduate School of Engineering, Utsunomiya University

³ KASAHARA Takehiko: Professor, Graduate School of Law, Toin University of Yokohama

している。サイバーコートは米国以外でもシンガポール、インド、韓国、英国など多くの国々で導入されている。

日本では、2004年に桐蔭横浜大学でサイバーコートのプロトタイプを用いた実証実験が行われた^{3,4)}。そこでは、裁判員システムへの有効性や、遠隔ビデオ会議システム等の実験が行われ、それぞれで有効性が証明された。また、内閣府の「未来投資戦略2017」では裁判手続きのIT化や裁判外紛争解決手続き(ADR)等の検討事項が記されており、今後の計画を2017年度内に決定し向こう5年間で実施していく予定となっている⁵⁾。これを受け、裁判手続き等のIT化検討会が平成29年10月30日に設置され、平成30年3月30日に、「裁判手続き等のIT化に向けた取りまとめ——「3つのe」の実現に向けて——」⁶⁾という提言が示された。ここでは、民事裁判におけるIT化として望まれる新しい姿と、その実現に伴う課題、実現プロセス等が示されており、サイバーコート実現に向けた一歩が踏み出された。

本論文では、サイバーコート実現に向けた重要な技術の一つである「プライバシー保護のための匿名化」について、それを自動化する機能のプロトタイプを、ニューラルネットワーク(NN)を用いて実装する際に必要となるNNのモデル構造について検討した結果を述べる。

II. 自動匿名化技術

1. 自動匿名化技術の必要性

前節では日本国憲法第32条が保証する「裁判を受ける権利」の解釈として伊藤正巳による「憲法上いかに裁判を受ける権利が全ての人間に公平に開かれており、形式的に正義へのアクセスの道が認められていても、事実上それを実行あらしめることを阻害する要因は少なくない。」を例として挙げた。この「形式的に正義へのアクセスの道が認められ

ていても」というのは現在の日本の裁判手続きが、まさに当てはまる。

これまでは、「事実上それを実行あらしめることを阻害する要因」として離島などの地方における「地理的要因」や、高齢者や身体障害者などにおける「身体的要因」が例として挙げられてきた。しかし、実際はそれだけでなく「法と裁判に関する知識的要因」が問題になる場合が多くある。

例えば、法や裁判に関しては全くの素人である国民が抱えている紛争を裁判で解決しようと考えた場合、ほとんどの場合「何もわからないため弁護士に相談しよう」と考えるであろう。これは決して悪いことではない。法に関するプロフェッショナルに助言してもらう方が解決までスムーズに進みやすい。しかしながら、「何もわからないから」つまり「他に方法がないから知識を得るために」相談しているとすればそれは、「事実上それを実行あらしめることを阻害する要因」が存在するからではないだろうか。(弁護士を否定しているわけでは無い事にご理解いただきたい。短時間では得られない「知識」「経験」「ディベートテクニック」の利活用や時間の節約のためには必要である。)

この「法と裁判に関する知識的要因」を解消するための方法の一つに、法と裁判の知識を持たない人々のための法と裁判に関する情報検索システムが考えられる。これは、インターネットを介して「自分自身で」これから行おうとしている紛争解決についての裁判手続きや類似判例の検索を可能とするものである。類似判例の検索については近年、日本でも研究が行われているが、実用段階のものはない。

この検索システムで重要となるのは蓄積されている判例の量である。裁判所のウェブサイトでは一部の判例のみの公開となっており、全件公開が望まれている。全件公開できない理由の一つにプライバシーの問題がある。プライバシー保護の観点から公開判例には匿名化が義務付けられているが、現在は手作業で

行うのが通例である。これは、現在の言語処理技術では文書の内容の解釈を十分に行うことができないためである。しかし、手作業では膨大なコスト（時間・料金）がかかるため全件公開を実現するためには匿名化処理の自動化は必須である。

また、未来投資戦略2017においても裁判手続きのIT化以外にもビッグデータの活用やそのためのデータ公開についても触れられており、裁判判例の全件公開は急務であると言える。

2. 自動匿名化の問題点とタスクの策定

匿名化処理（以下、伏字化処理）の例を図1に示す。可読性確保のため必要最低限の伏字化に留める必要がある、また、固有名詞辞書を用いた手法ではその構築の困難さと、判例内でマッチングしたとしてもその単語が他の意味で使用されている（例えば「正月」は「むつき」という苗字である場合がある。）可能性があるため現在では手作業により処理されている。全ての裁判データを公開することを考えると、手作業では到底追いつける量ではない。

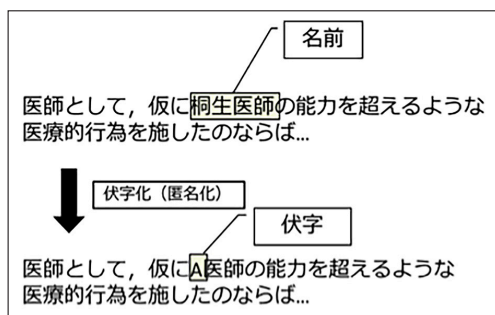


図1 伏字化処理の例

そこで我々は、これまでに手作業で伏字化された判例を真のデータとして学習データに用い、機械学習することで未だ伏字化されていない文書を入力すると自動で伏字化してくれるシステムを最終的な目標として、研究を開始した（図2）。

また、本研究では近年自然言語処理で大きな成果を上げているニューラルネットワーク

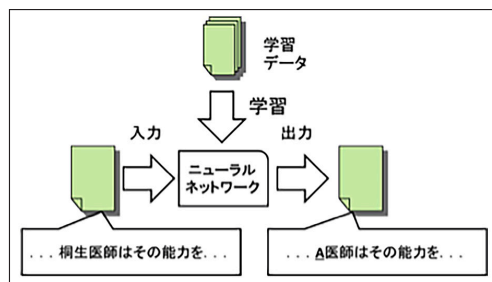


図2 伏字化システムの概要

を用いるものとし、その具体的なタスクを「ターゲット単語を伏せるべきかどうかを、その周りの単語を入力として判断する」と策定した。

このタスクの概要を図3に示す。図3はターゲット単語（Dr. Kiryu）が伏字にすべきかどうかを予測するために、周辺単語を入力としている様子を示している。

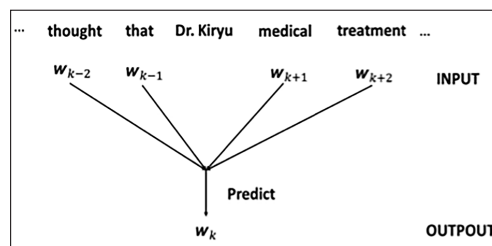


図3 予測タスクの入出力

尚、内部の計算方法に関しては次章以降で詳細に述べる。

図3では、入力は前後2単語となっているが、これは目的に応じてパラメータとして調節する値である。モデルの定義の時点で再度紹介するが、このパラメータのことをウィンドウサイズと呼ぶ。

尚、本論文ではパラメータ最適化については論ぜず、ニューラルネットワークの他の論文を参考にし、前後5又は10を利用することとした。

また、出力の予測結果は「伏字にするかどうかの二択」ではなく、「ターゲット単語は何か？」というターゲット単語を当てるタスクとした。なぜなら、このタスクは一般的な文書を利用した場合の先行研究が多くあり、参考にできるモデルが豊富であるためである。

このタスクを学習する場合、伏字を表す文字を定義しておく必要があるが、ここでは、伏字は「A」で表すこととした。

Ⅲ. 自然言語処理のためのニューラルネットワークモデル

1. 既存ニューラルネットワークモデル (1) Recurrent Neural Network (RNN)

RNN は連続データを処理する仕組みであるが、**図 4** の例では予測したい時系列の5つ前のデータからしか考慮していない。

この値は、パラメータとして、モデル構築者が自由に選ぶことができる。

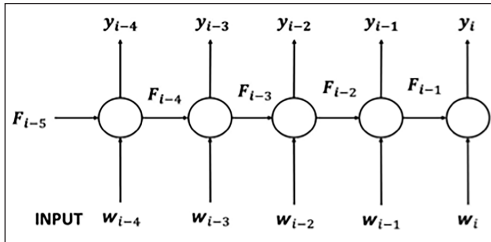


図 4 RNN の概要～入力は 5 つ前まで～

この例では、 $w_{i-4}, w_{i-3}, w_{i-2}, w_{i-1}, w_i$ を入力として出力を決定している。

言語処理では、 w_{i+1} を入力として、その次に来る単語を予測することがよくある。この使用例が本研究の伏字位置予測タスクに利用できると思った。尚、本論文では、RNN の拡張である LSTM (Long Short Term Memory) モデルを利用した。

このモデルは 2010 年に Mikolov らが提案したものであり、人間の脳の長期記憶と短期記憶を模倣したモデルである⁷⁾。

RNN の問題点の一つに勾配消失問題と言うものが存在する。勾配消失問題とは多層のニューラルネットワークにおいて、誤差逆伝搬の際に伝搬する誤差が段々と消失してしまう問題である。このため多層になるほど学習速度が低下し、学習に莫大な時間がかかることになる。LSTM では、それが改善されている。先ほど説明したように、RNN は連続

データを扱うことができ、いくつ前までのデータを扱うのかはモデル構築者がパラメータとして扱うことができる。昔のデータまで振り返るため、ニューラルネットワークの層が増えていくこととなる。

したがって、この問題を解決した LSTM モデルを使用することが妥当である。

また、LSTM は入力ゲート、出力ゲート、忘却ゲートなどの仕組みを用いることで長期記憶を扱うことが可能である。

(2) CBOW モデル～word2vec～

次に、ニューラルネットワークモデルの一つである CBOW モデルに説明する。

word2vec はニューラルネットワークにより単語の概念を低次元ベクトルとして取得できるツールであり、ニューラルネットワークのモデルである CBOW (Continuous Bah-of-Words) と Skip-gram の総称である。近年、word2vec は司法試験のテキスト推論、法的文書解析、法的 QA システムなどでも使用されている。

CBOW は注目単語の前後数単語 (デフォルトは 5. この数を決定するのが window オプション) の 1-of-K ベクトルを足し合わせた物を入力として、注目単語を出力するようなニューラルネットワークを学習するモデルである (**図 5**)。

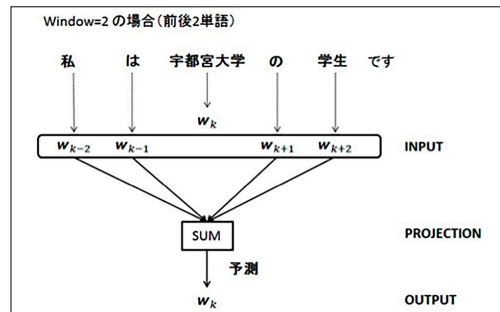


図 5 CBOW

一方 Skip-gram はある単語が入力されたとき、その周辺の数単語 (window により指定) 中の 1 単語を予想していくという学習モデルである (**図 6**)。

Mikolov らによると、CBOW モデルは

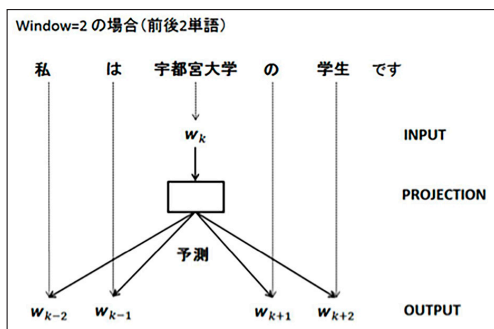


図 6 Skip-gram

NNLM (Neural Network Language Mode) よりも構文的、意味的な処理とも良い結果であった。

さらに、Skip-gram は CBOW モデルよりも構造的な処理は少々劣るが (しかしながら、それでも NNLM よりもまだ良い結果であった)、意味的な処理は他のモデルよりも圧倒的に良い結果となった。

このことを考慮すると、Skip-gram モデルが精度の面で優秀であり使用すべきであると考えられるが、II. 2で策定されたタスク「ターゲット単語を伏せるべきかどうかを、その周りの単語を入力として判断する」との相性を考えると CBOW モデルのほうが相性が良い。

したがって、本研究では CBOW モデルを用いることとした。

2. CBOW の単語予測への応用事例

CBOW モデルの用いた単語予測タスクに関する既存研究を紹介する⁷⁾。以下では、CBOW モデルの「言い換え能力」に焦点をあて単語予測タスクへの応用を検討したものである。実験では通常の CBOW モデルに加え、ターゲット単語の左右を分割した LR モデル、NNLM のようにすべての入力を分割した WO モデルを使用した。両モデルとも予測問題に関しては CBOW モデルよりも良い精度が出ることが示されており、CBOW が持っている言い換え能力の有効性についても考察されていた。

今回の伏字位置予測のタスクについてもこ

の言い換え能力に着目し、CBOW モデルの構造を取り入れたモデルを構築する。

IV. 伏字予測のためのニューラルネットワークモデルの定義

本節では前節までで行った有効な CBOW モデル構造や既存の LSTM の言語モデルへの適用、人間による手作業での伏字化の模倣を行い、伏字予測のためのニューラルネットワークモデルを定義していく。

1. 既存の LSTM モデル

オーソドックスな LSTM モデルを図 7 に示す。基本モデルであるため、入力として使用するのはターゲット単語よりも前に出現した単語でウィンドウサイズは 5 とする。

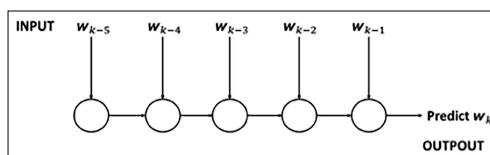


図 7 Base Method LSTM

2. 手法 1: 人間による作業を模倣したモデル

提案手法 1 として、「人間による作業を模倣したモデル」を定義する (図 8)。これは、人間が伏字化を行うときの作業を模倣したモデルで、ターゲット単語を挟み込むように前後から解析していくモデルである。通常の LSTM の言語モデルであれば文章を前から後ろの順で読み込ませていくが、本手法ではターゲット単語の後ろ側の入力は読み込ませ

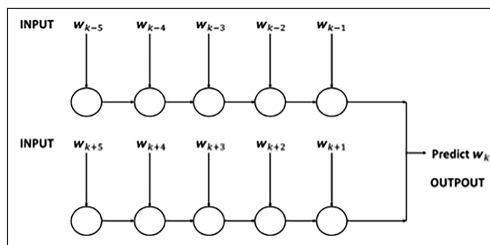


図 8 Method 1

る順が逆となっている。これは、LSTMがいくら長期記憶を扱えるようにしたモデルと言えども直前に読み込まれた入力のほうが影響を与えやすいため、最も伏字かどうかの判断に必要となるターゲット単語の隣接単語を最後に読み込ませるように配慮したためである。

3. 手法2：CBOWの構造を取り入れたモデル

提案手法2として、「CBOWの構造を取り入れたモデル」を定義する(図9)。これは8.6節の実験で分かった「伏字予測には足し合わせベクトルが有効に働く」という結果を利用したモデルである。既存のLSTMモデルの入力を全て足しあわせ、最終層への出力の際にLSTMの出力と連結し、そこに重みとバイアスを掛けて出力するモデルとなっている。

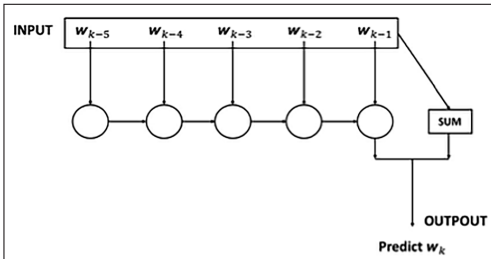


図9 Method 2

4. 手法3：手法1と2を組み合わせたモデル

提案手法3として、「手法1と手法2を組み合わせたモデル」を定義する(図10)。手法2ではベクトルの足し合わせがターゲット単語より以前に出現した単語のものであ

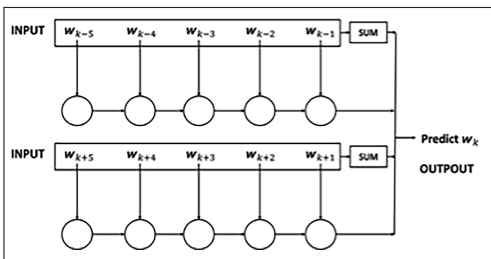


図10 Method 3

たが、手法1と組み合わせることで後ろ側も計算することになる。つまり、LRモデルを取り込んだようなモデルとなっている。

V. 伏字位置予測実験

1. 実験諸元

今回使用したコーパスは1993年から現在までの裁判判例約60,000件の中から、ランダムに学習データ50,000件を抽出し、残りからランダムに10,000件のテストデータを抽出したものを使用した。裁判判例内の伏字位置予測にストップワード(句読点、助詞に隣接して出現する)は必要であると考えられるため、ストップワードの除去は行っていない。ただし、文の流れを切らないようにするためカギ括弧とその中身は除去してある。また、判例内で使用されている各伏せ字を表す文字(大文字アルファベット, 甲, 乙等)は全て「A」に変換した。モデルのウィンドウサイズについては、手法2はターゲット単語の前5個のみ、手法1,3はターゲット単語の前後5単語となっている。LSTMの隠れ層のサイズは100、学習率は0.001で、埋め込みベクトルのサイズは200、バッチサイズは200とした。辞書に登録する最小の出現回数は0としている。評価方法としては、ニューラルネットワークを利用した予測問題で良く使われるppl(perplexity)を採用した。ニューラルネットワークを用いた単語予測の既存研究で用いられ続ける評価指標で、正解語と答える確率の逆数を表す。つまり、「解答を何択まで絞ることができたのか」を表す指標とも言える。pplを求める式を以下に示す:

$$ppl = \frac{1}{\text{probability}(\text{correct word})}$$

pplは何択まで絞ることができたのかを表す指標なので最も良いスコアが1である。このスコアが大きければ大きいほど悪いスコアであると言える。今回の実験では、テストデ

一タの全ての単語に対して予測タスクを解かせその平均を計算する ppl と、テストデータの中で解答が伏字、すなわち「A」となる問題のみを解かせその平均をとったものを upw_ppl とし、この2つのスコアについて計算した。

2. 実験結果

実験の結果を表 1 に示す。

表 1 伏字位置予測の実験結果

	LSTM	Method1	Method2	Method3
ppl	4.851	4.656	4.603	4.462
upw_ppl	56.277	37.343	72.463	77.031

まず、ppl の結果を見ると手法 1 のスコアはシンプルな LSTM のスコアよりも 0.195 良い結果となっており、手法 2 はシンプルな LSTM よりも 0.247 良いスコアとなった。手法 2 は手法 1 よりも 0.053 だけ良い結果となっており、入力ベクトルの足し合わせ (CBOW の構造) の有効さが目立つ結果となった。また、両手法を組み合わせた手法 3 は手法 2 よりもさらに 0.141 だけ良い結果となっており、シンプルな LSTM モデルよりも 0.389 良くなっている。これらの結果から、両手法とも期待通りの向上を見せていることが分かる。次に、upw_ppl のスコアに注目する。ppl のとスコア差が少なくとも 32.492 あることから、他の語の予測に比べて伏字を予測するタスクが難しいことが分かる。各種法について比較していくと最も良いスコアとなったのが、手法 1 でありシンプルな LSTM モデルに比べて 18.934 だけ良くなっている。しかしながら、手法 2 ではシンプルな LSTM モデルよりも 16.186 だけ悪くなっていることが分かる。また、手法 3 では LSTM モデルよりも 20.745 悪い結果となった。結果をまとめると、ppl に関してはいずれの手法ともに精度向上が見られ、特に手法 2 がもっとも良い。逆に upw_ppl に関しては、手法 1 のみ精度が向上したが、手法 2、3 で

は悪化する結果となった。

3. 考察

実験結果から、ppl では手法 2 が精度向上を果たしたのに対し、upw_ppl に関しては精度低下を招いた理由を考察していく。まず、ppl の精度が向上した理由に関しては、今まで以上に埋め込みベクトルの学習が上手く行ったことが考えられる。入力ベクトルに伏字などの捉えにくい単語が入った場合などは、埋め込みベクトルの学習を強化していない既存手法では上手く予測することが困難であった。しかしながら、CBOW モデルの足し合わせ構造を導入することで、より質の良い埋め込みベクトルが構築でき、言い換えや近い単語を認識できるようになった。このことから、ppl の精度が手法 2 によって向上したと推測される。

次に、upw_ppl において手法 2 が有効でないという結果について考察する。まず、精度が悪化した理由であるが、これは CBOW の「言い換え能力」が悪い方向に作用したのだと考える。「伏字になるべき単語」はもと、「原告」、「被告」、「医者」、「父」、「息子」、「法人」など様々なロールを持つ概念である事がほとんどである。手法 2 により質の良い埋め込みベクトルを得て、単語の言い換え能力を得たモデルは最終的な予測の場面で選択肢を絞りきれないという問題点が出てきてしまったと考えられる。しかしながら、ppl ではなぜ予測の場面で複数の選択肢が出てこないのかという疑問が出てくる。ppl では伏字以外にも判例中出现する全ての単語を対象としてテストをしていることを考慮すると、伏字と予測されるべき単語はたくさんの言い換え語を持っているが、その他の単語はそれほど言い換え語を持っていないと考えられる。例えば、助詞、形容詞、形容動詞、副詞など。名詞であっても伏字となる語ほどの言い換え語を持つ単語は少ない。このため、手法 2 が upw_ppl のみに悪影響を及ぼしたと推測される。また、手法 3 が手法 2 と同様



図 11 伏字位置予測ツールのプロトタイプ画面

の結果となった理由も同様に、手法2の言い換え能力を手法1の構造がより強化してしまったのだと考えられる。

手法2は upw_ppl のスコアが悪くなってしまったが、この結果から一概に手法2は悪手であるとは言いきることはできない。予測問題として考えてみると、言い換え語を出力してしまったとしてもそれは間違いとは言いきれない。文脈に当てはまるように予測できているのなら正解としても良いのではないかと考えられる。今回の upw_ppl スコアが悪い事がモデルの能力を決定するものとは不適切と思われる。しかしながら、今回の実験はテストデータを 10,000 件準備しているため、手作業で確認するのは現実的に難しい。例えば、埋め込みベクトル等を使い、伏字を表す文字の埋め込みベクトルからの距離とその単語の確率の積を取るなどの工夫が必要になると思われる。CBOW を拡張したモデルで大量のデータに対して予測問題を行った研究は無く、今後の検討課題としたい。

VI. おわりに

1. まとめ

現在、判決文を公開する際には、プライバシーを保護するために、公開してはならない語の匿名化、すなわち伏字化する作業を行っている。この作業は、文脈の認識を伴う作業であり、文中での表記ゆれ（苗字が最初だけにしかついていないなど）に対応するため、人手により行われてきた。今後より多くの判決文を公開するためには、この作業をより高速化する必要があり自動化が必須となる。本論文では自動伏字化のためにニューラルネットを用いてプロトタイプを作り、評価を行った結果について述べた。質の良い埋め込みベクトルを得ることができ、CBOW モデルと連続データを効率良く扱うことのできる LSTM を参考にし、新たなモデルを定義し実験した。その結果、CBOW モデルを参考にしたモデルでは、最終的に伏字かどうかを決定する場面で言い換え語がたくさん候補として挙がってしまい、良い精度を得ることが

できなかった。今後は、言い換え語に関する定量評価方法を考え直すとともにモデルの拡張を行っていく必要がある。

2. 今後の予定

今後は今回の実験で生じた「言い換え語」に関する問題に注意しながら伏字位置予測の精度を上げていく必要がある。

また今後は、伏字位置のチューニングのためには、法学領域の専門家の方々の協力が必須であるため、より分かりやすい GUI ツールを構築したいと考えている。現在制作途中の GUI ツールの画面を図 11 に示す。アンダーラインは固有名詞辞書と形態素解析器 MeCab による予測であり、あみかけが提案手法である。ニューラルネットワークによる手法はターゲット単語がどのような文字であるかは考慮しないので、裁判用語や一般的に伏せるわけがない単語もヒットしてしまっている。今後はこのような単語を法令用語辞書等を利用して事前に登録しておき、前処理段階で除外しておく等の対策により、伏字位置予測の精度を向上させることができると考えている。

脱稿後の 2018 年 3 月 30 日、「裁判手続等の IT 化に向けた取りまとめ」が出され、裁判の IT 化を推進することとなった。

- 6) 裁判手続等の IT 化検討会, 「裁判手続等の IT 化に向けた取りまとめ——「3つのe」の実現に向けて——」, <https://www.kantei.go.jp/jp/singi/keizaisaisei/saiban/dai8/siryou.pdf> (2018).
- 7) Mikolov, Tomas, et al. “Efficient estimation of word representations in vector space.” arXivpreprint arXiv:1301.3781 (2013).

【注】

- 1) <http://www.taf.or.jp/report/23/index-1/page/p069.pdf> [Mar, 25, 2018]
- 2) <http://law.wm.edu/academics/intellecualife/researchcenters/clct/> [Mar, 25, 2018]
- 3) Takehiko Kasahara, “Cybercourt - Court Technology”, Journal of the Japanese Society for Artificial Intelligence, vol. 23, No. 4, pp 513-, Jul. 2008]
- 4) Takehiko Kasahara, “Court Technology for Civil Procedure”, Festschrift for seventieth birthday of Prof. Takeshi Kojima II , pp. 961-, Sep. 2009
- 5) http://www5.cao.go.jp/keizai-shimon/kaigi/minutes/2017/0609/shiryo_07.pdf.