

特許文献への分類付与と付与根拠の推定

Classifying patent document and Suggesting grounds for assigning the patent classification

特許庁 審査第一部調整課企画調査官

近藤 裕之

1999年入庁。特許審査、審判に従事のほか、情報技術統括室で特許庁情報セキュリティポリシー策定、システム計画などを担当すると共に、内閣官房で政府の政策の総合調整、一般財団法人工業所有権協力センター（IPCC）を経て、2018年4月より現職。

1 はじめに

平成26年に閣議決定された『「日本再興戦略」改訂2014』及び知的財産戦略本部決定された「知的財産推進計画」において、「世界最速・最高品質の特許審査」の実現を目指すこととされ、その実現のためには内外国の先行技術文献調査を効率的に実施するための環境整備が重要である。

特許庁では、平成28年度から、人工知能（AI）技術の活用に向けた検討を行い、同年に行った調査研究（「人工知能技術を活用した特許行政事務の高度化・効率化実証的研究事業」）の結果を踏まえ、平成29年4月に、「特許庁における人工知能技術の活用」についてのアクションプランを公表した。

このアクションプランに沿って、平成29年度は、「電話等の質問対応」、「紙出願の電子化」、「特許分類付与（テキストに基づく付与）」、「先行技術調査（検索式の用語の拡張、ヒット箇所のハイライト表示）」、「先行図形商標の調査」、「指定商品・役務調査」の6つで調査事業を実施しているが、本稿では、「特許分類付与（テキストに基づく付与）」の調査事業について紹介する。

特許文献への分類付与に関して、特許庁では、平成28年度に外国特許文献への分類付与に関する機械学習の活用可能性調査を実施した。この調査において、特許分類を付与するための根拠箇所の記載が、特許文献の

ごく一部の記載箇所にしか現れない分類や、逆に高頻度で使用されている単語の組合せによって記載されている分類など、用語の使用頻度が一定の範囲にある場合でなければ、ノイズとなってしまい、付与精度が低下する原因になっていることが挙げられている。そこで、平成29年度の調査事業においては、特許分類（Fターム）や、その分類に対する明細書中の付与根拠箇所（段落単位）の推定について検討を行うこととした。

なお、本稿は、平成29年度に実施した調査事業の結果を踏まえ、著者らの私見に基づいて記載したものであり、特許庁としての意見・見解を表明するものではない。

2 機械学習モデルごとの推定精度検証

本調査事業においては、各種の機械学習モデルを用い、特許文献に付与すべき特許分類（Fターム）や、その分類に対する明細書中の付与根拠箇所（段落）を機械推定し、その精度を比較評価した。また、検証手法自体も、[検証1] 学習データに正解データを用いずに機械学習させて構築した機械学習モデルを用い、特許分類（Fターム）に対する付与根拠箇所推定精度を評価する検証、及び[検証2] 学習データに正解データを用いて機械学習させて構築した機械学習モデルを用い、特許分類（Fターム）に対する付与根拠箇所推定精度を評価し、さらにFターム自体の付与精度も評価する検証である。

本章では、本調査事業で実施した上記2つの検証について、用いた機械学習モデル及び推定精度評価結果の概要並びに得られた知見等を紹介する。

2.1 [検証1] 機械学習モデル作成に正解データを用いない手法

本検証では、人手で作成した付与根拠箇所の正解データについては学習に用いずにFタームの付与根拠箇所を推定可能な4種類の機械学習モデル（NAM、CNN-NAM、mi-SVM、mi-NN）を構築し、各機械学習モデルにおけるFタームの付与根拠箇所推定精度を比較評価した。

(1) 機械学習モデルの概要

検証1で用いた、4種類の機械学習モデルの概要について、それぞれ紹介する。

① NAM（ニューラルアテンションモデル）^{3, 4, 5}

文献に現れる各単語の注目度合（Attention）を学習し、分類付与に貢献する単語と、そうでない単語を識別した上で分類を付与するニューラルネットワークである。NAMは、単語ごとに文脈を考慮した表現を抽出するための「連続的な要素に対するニューラルネットワーク（NN）」と、文脈を考慮した表現をもとに、各単語の注目度合を表す値を得るための「注目度合を決定するAttention層」から構成される。

このNAMの特徴は、「(a) 類似する文書を検索する」と、「(b) 類似する検索による文書である」という単語の要素（類似、文書、検索という単語で構成）は同じだが、意味が全く異なる両文を区別して学習できる。なお、このような連続的な要素を区別できるニューラルネットワークとしては、リカレントニューラルネットワーク（RNN）、畳み込みニューラルネットワーク（CNN）がよく知られる。

Fタームに対する付与根拠箇所推定にNAMを適用した場合の構成を図1に示す。

NAMでは、どの記載箇所注目すると最終的なFターム付与精度が向上するのかわかるという情報を元に、注目の仕方を学習する。付与されているFターム情報に基づいて最適値を自己学習するので、過去の特許文献に付

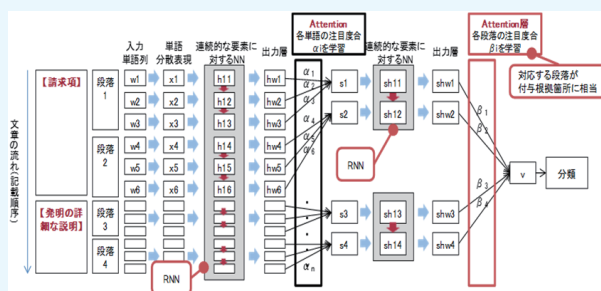


図1 NAMによる付与根拠箇所推定の構成

与されたFタームに対する付与根拠箇所を教師データとして使用する必要がない。

② CNN-NAM（畳み込みニューラルアテンションモデル）

このモデルは、上記①で述べたNAMの隠れ層（RNN）を、CNN（畳み込みニューラルネットワーク）に置き換えたニューラルネットワークであり、NAMと比較し、並列処理によって学習を高速化でき、マシンリソースも削減できる特長がある機械学習モデルである。

Fタームに対する付与根拠箇所推定にCNN-NAMを適用した場合の構成を図2に示す。

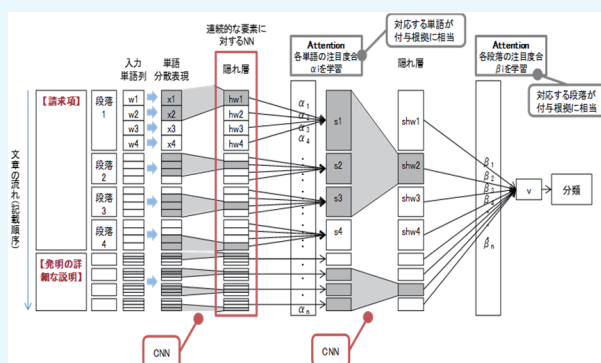


図2 CNN-NAMによる付与根拠箇所推定の構成

上記①で述べたNAMの隠れ層で採用しているRNN（LSTM、GRU等）の代わりに、CNNの畳み込みを施すことにより、学習精度を低下させずに学習を高速化させる。

③ mi-SVM（マルチインスタンスサポートベクターマシン）^{6, 7}

このモデルは、複数の段落から構成される特許文献に対して、各段落が付与根拠箇所とみなされるかをFタームごとに推定し、付与根拠箇所とみなされない段落を除

外しては再学習するという処理を繰り返すことにより、付与根拠箇所を絞り込んでいく機械学習モデルである。

mi-SVM は、まず、ある F タームが付与された文献（正例文献）の全段落を、付与根拠箇所である（正例段落）とし、それ以外の文献（負例文献）の全段落を付与根拠箇所でない（負例段落）として初期設定する。次に、これらの段落を SVM で初期学習し、学習器を生成する（①初期学習）。そして、生成された学習器を使用して、教師文献自身に対して、各段落が付与根拠箇所であるかを推定する（②推定）。その結果、付与根拠箇所でないとして推定された段落を正例落から除外し、更新された正例段落及び負例段落を使用して再度 SVM で学習し、学習器を生成する（③再学習）。このように、正例段落を更新しては学習・推定するという処理を繰り返すことにより、付与根拠箇所（正例段落）を絞り込んでいく。

このモデルを F タームに対する付与根拠箇所推定に適用した場合の構成を図 3 に示す。

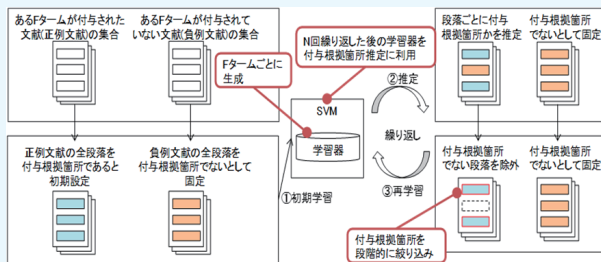


図 3 mi-SVM による付与根拠箇所推定の構成

④ mi-NN (マルチインスタンスニューラルネットワークモデル)

このモデルは、mi-SVM において使用する SVM の代わりに、ニューラルネットワークモデルである fastText^{8,9} を適用した機械学習モデルである。単語の分散表現を事前学習する際に、単語を構成する部分文字列の分散表現を加味することにより、表現力の高い単語分散表現を生成している点、また、fastText は単語の頻度を活用し、不必要な計算を減らしながら学習するため、他の事前学習手法と比べて学習に要する時間が短い点が特長である。

このモデルを F タームに対する付与根拠箇所推定に適用した場合の構成を、図 4 に示す。

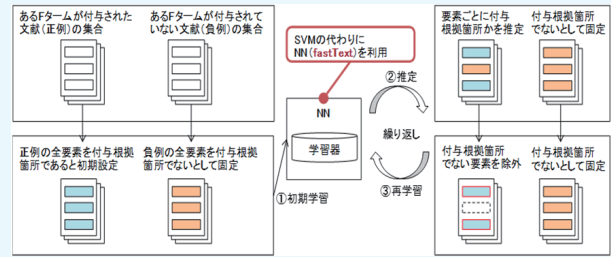


図 4 mi-NN による付与根拠箇所推定の構成

(2) 検証 1 における機械推定精度評価結果

本検証では、12 の技術分野（テーマ）について、一つの技術分野（テーマ）あたり 20 個（合計 240 個）の F タームを選定し、日本語の特許文献を対象とし、各テーマ 200 文献の評価データ（正解データ）に対して付与根拠箇所の推定精度評価を行った。推定精度評価の指標としては、個々の F タームに対する F 値を評価文献ごとに算出し、算出結果を平均する「マクロ平均 F 値」と、個々の F タームに対する個々の評価文献ごとの結果をマージ（加算）してから F 値を算出する「マイクロ平均 F 値」を用いた。マクロ平均は、正解付与根拠箇所の数が少ない F ターム（または文献）の精度の振る舞いに影響されやすく、一方、マイクロ平均は、正解付与根拠箇所の数が多い F ターム（または文献）の付与精度の振る舞いに影響されやすい指標である。さらに、推定精度評価の指標として、「抽出率」も採用した。抽出率とは、「評価文献を構成する総段落数に占める、付与根拠として機械が推定した段落数の割合」である。人手の正解データと機械推定した付与根拠箇所の抽出率とを比較することで、機械がどれだけ付与根拠箇所を絞り込めているか（あるいは、過度に絞り込んでいないか）の評価が可能である。

検証 1 において、付与根拠箇所の推定精度は、マイクロ平均 F 値では、mi-SVM、NAM、CNN-NAM、mi-NN の順に高く、一方、マクロ平均 F 値では、NAM、mi-SVM、CNN-NAM、mi-NN の順に高いという結果が得られた。NAM は、マイクロ平均 F 値では mi-SVM に及ばないが、マクロ平均では上回っていることから、mi-SVM よりも、比較的安定した出力結果を出せていると考えることができる。

また、付与根拠箇所の抽出率を比較した結果、人手で付与した正解付与根拠箇所の抽出率に対し、mi-SVM、

mi-NN の抽出率は比較的近いが、NAM、CNN-NAM ではかなり高くなっており、付与根拠箇所を適切に絞り込めていないことが分かった。

2.2 【検証 2】 学習データに正解データを用いた手法

本検証では、検証 1 で利用した特許文献のテキストデータ、及び、特許文献に付与された F タームに加えて、人手で付与した付与根拠箇所情報の正解データも学習データとして利用する 3 種類の機械学習モデル (NAM+、mi-SVM+、mt-NAM) を構築し、当該モデルを用いた付与根拠箇所推定、及び、F タームを付与した際の付与精度評価を行った。なお、F ターム付与の精度評価には、ベースラインの機械学習モデルとして SVM を採用した。

(1) 機械学習モデル

本節では、検証 2 で用いた、4 種類の機械学習モデルの概要について、それぞれ紹介する。

① SVM (サポートベクターマシン)

これは、最も広く適用されている機械学習モデルの一つであり、本検証では、F ターム付与精度評価のベースラインとして採用した。SVM は、基本的には、2 クラスを識別する機械学習モデル (2 値分類モデル) であり、2 クラスを識別する学習モデルを教師文書から生成し、新規文書に対してどちらのクラスに属するかを識別する。例えば、受信した電子メールがスパムメールであるか否かという 2 クラスを想定した場合、SVM はスパムメールである電子メール群 (正例) と、スパムメールでない電子メール群 (負例) を教師文書として学習し、学習モデルを生成する。そして、新規に受信した電子メールに対して、スパムメールであるか否かを識別し、識別結果を出力する。

SVM の最大の特徴は「マージン最大化」と呼ばれる考え方にある。一般に 2 クラスの境界を学習する場合、図 5(a) に示すようにいろいろな教会の引き方が存在する (図では、理解促進のために、2 次元空間での境界を示しているが、実際には n 次元空間での境界となる。) これらの境界のうち、SVM では、図 5(b) に示すように、

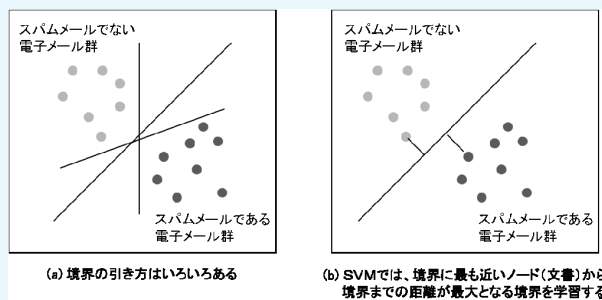


図 5 SVM の特長 (マージン最大化)

境界に最も近いノード (文書) から境界までの距離が最大となるように境界を学習する。この考え方が「マージン最大化」である。このマージン最大化により、新規文書に対する識別精度を向上させる (汎化性能を高める) ことができる。

F ターム付与は、一つの特許文献に一つ以上の F タームが付与されるマルチレベル分類である。したがって、マルチレベル分類のタスクに対して、2 値分類モデルを適用することになる。SVM をマルチレベル分類に適用する方式としては、One-vs-rest、One-vs-one という 2 種類の方式がある。本事業では、学習モデルの作成時間を考慮し、One-vs-rest 方式を採用した。その構成は、図 6 に示す。

One-vs-rest 方式は、図 6 に示すように、個々の分類に着目して、その分類を付与すべきか否かを判定する学習モデル (2 値分類モデル) を生成し、付与すべきと判定された分類をすべて出力する方式である。したがって、生成される学習モデルは、付与対象とする F タームの数だけ生成される。本方式では、ある F タームを付与すべきかを判定する際に、他の分類の影響を受けない。

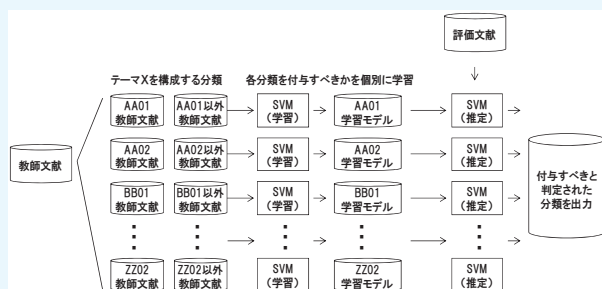


図 6 SVM による F ターム付与方式 : One-vs-rest 方式

なお、SVM は、付与根拠箇所を活用せずに F ターム付与した場合の付与精度を比較検証するためのベースラインモデルである。そのため、付与根拠箇所を推定する

ことはせずに、Fタームのみを付与して出力することとした。

② NAM+ (ニューラルアテンションモデル+)

このモデルは、検証1で述べたNAMを、「付与根拠箇所推定ネットワーク」及び「推定された付与根拠箇所を使用したFターム付与ネットワーク」の2つに分割したモデルである。NAMを二つのネットワークに分けることによって、前者の付与根拠箇所推定ネットワークで付与根拠箇所を教師データとして付与根拠箇所の推定方法を学習し、後者のFターム付与ネットワークでFタームを教師データとしてFタームを付与するかを学習できるようになる。

付与根拠箇所推定ネットワークは、Attentionと同様の方法で付与根拠箇所を推定し、推定した付与根拠箇所と教師データの付与根拠箇所を突き合わせることで付与根拠箇所の推定方法を学習する。

Fターム付与ネットワークは、推定された付与根拠箇所を使って、予測に使うデータを絞り込んだ上で、上記したNAMと同じ方法でFタームを付与する。

なお、本モデルは、上記したCNN-NAMで置き換えることによって、CNNを使ったモデルとしても実装できる。

NAM+を付与根拠箇所推定及びFターム付与に適用した場合の構成を、図7に示す。

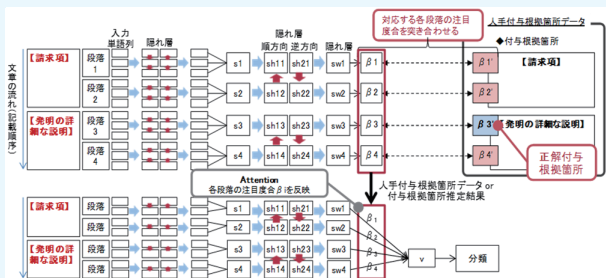


図7 NAM+による付与根拠箇所推定及びFターム付与の構成

③ mi-SVM+ (マルチインスタンスサポートベクターマシン+)

このモデルは、mi-SVMを拡張した機械学習モデルである。具体的には、mi-SVMに対して2点の拡張を施している。1つ目は、mi-SVMでは、付与根拠箇所が未知である文献を教師文献として学習するのに対し、

mi-SVM+では、図8に示すように、付与根拠箇所が既知の文献(機械または人手によって特定された付与根拠箇所を持つ文献)を教師文献として使用している。付与根拠箇所が既知の文献を学習に使用することにより、付与根拠箇所の推定精度及びFタームの付与精度の向上が期待できる。2つ目は、図8に示すように、入力文献に対して、まず、付与根拠箇所推定用のSVMによって付与根拠箇所を推定する。そして、出力された付与根拠箇所に含まれる単語から素性データを抽出し、元の評価文献から抽出される素性データに追加(結合)して、Fターム付与用のSVMによってFタームを付与する。

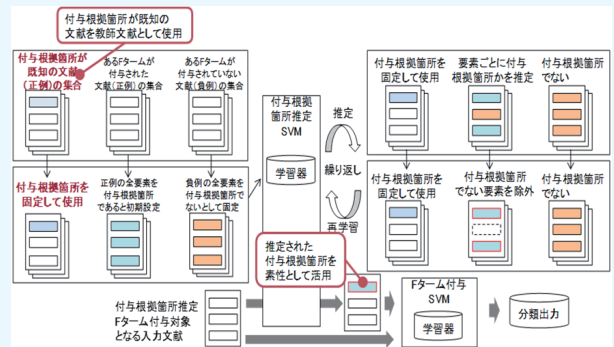


図8 mi-SVM+による付与根拠箇所推定及びFターム付与の構成

④ mt-NAM (マルチタスクニューラルアテンションモデル)

このモデルは、付与根拠箇所推定とFターム付与という二つのタスクを、同じモデル上で学習するモデルである。

このモデルでは、まず、上記したNAMと同様の手順で注目度合い β によって、Fターム付与の予測値を得る。そして、教師データの教師信号(付与Fターム)と予測が近くなるようにモデルパラメータを最適化する。

次に、NAMと同様に、段落単位の注目度合い β を学習し、注目度合いの高い段落を付与根拠箇所の予測値として出力する。そして、学習データの教師信号と予測が近くなるように追加で学習を行う。この手順は、Fターム付与に関して学習が完了したネットワークを、付与根拠箇所を適切に出力できるように微調整することに相当する。

このモデルの特徴は、上記したNAM+は、付与根拠箇所推定、Fターム付与のそれぞれに対して異なるネッ

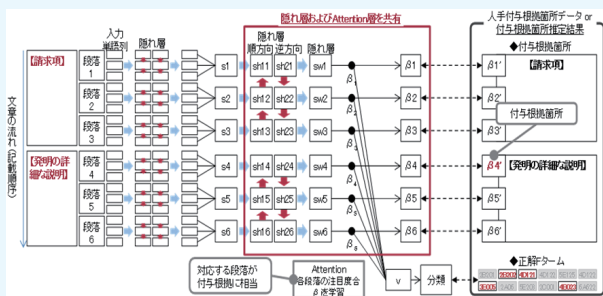


図9 mt-NAMによる付与根拠箇所推定及びFターム付与の構成

ネットワークを持っているのに対し、mt-NAMは、付与根拠箇所推定とFターム付与という二つのタスクを同じモデル上で学習し、隠れ層及びAttention層を共有するので、学習の相乗効果により精度が向上する点にある。

(2) 検証2における機械推定精度評価結果

本検証では、検証1と同一の技術分野（テーマ）、Fターム、日本語特許文献を対象とし、各テーマ200文献の評価データ（正解データ）に対し、人手によって特定された付与根拠箇所正解データを各機械学習モデル（NAM+、mi-SVM+、mt-NAM）で学習した場合の、Fタームの付与根拠箇所推定精度及びFターム付与精度を比較評価した。各機械学習モデルの付与根拠箇所推定精度を比較した結果、マイクロ平均F値及びマクロ平均F値の両者とも、mt-NAM、NAM+、mi-SVM+の順で精度が高いことが分かった。また、本検証では、Fタームの付与精度についても算出し、各機械学習モデル間で比較評価したところ、NAM+、mt-NAM、SVM、mi-SVM+の順で精度が高くなる結果が得られた。

上述のとおり、今回検証したモデルのうち、付与根拠箇所の推定精度が最も高いモデルはmt-NAMであり、他方、Fターム付与精度が最も高いモデルはNAM+であった。mt-NAMでは、まずFターム付与を学習し、次にその学習結果を付与根拠箇所に適用してモデルパラメータを調整するが、本検証では人手で作成した高精度の付与根拠箇所データを学習に利用したため、この調整が功を奏して付与根拠箇所推定精度が向上したと考えられる。一方、NAM+では、付与根拠箇所の学習結果をFターム付与に適用するため、Fターム付与精度が向上したと考えられる。さらに、付与根拠箇所の抽出率を比較した結果、mi-SVM+、mt-NAM、NAM+、の順で低くなっており、mi-SVM+、mt-NAMにおいては、正解

付与根拠箇所の抽出率を下回っていることが分かった。

2.3 推定精度向上に向けた考察

本調査事業における検証では、最新の機械学習技術を用い、特許分類（Fターム）に対する明細書中の関連箇所（付与根拠箇所）の推定精度を調査したが、各検証を通して、さらなる推定精度向上化のための課題やその対策案等について知見が得られた。本節ではそれら得られた知見について紹介する。

(1) 複合語の分割

本検証では、形態素解析によって特許文献テキストを単語に分割している。その際、分割された単語（名詞、動詞）だけでなく、連続する二つの名詞を一つの複合語（例えば「情報」と「検索」を「情報検索」という一単語とする）として追加することにより、付与根拠箇所推定精度が向上する可能性がある。

(2) 機械推定の閾値の適切な設定

検証1のNAM及びCNN-NAMの抽出率は、人手で付与した正解付与根拠箇所の抽出率に対し、大幅に高い値となっていた。これは、付与根拠箇所であるかを判定する閾値が固定されており、Fタームによっては絞り込みが適切に機能していなかったことに起因すると考えられる。最良となる閾値を探索することで、抽出率を低くし、より妥当なF値を得られる可能性がある。

(3) 特許文献の内容・傾向を考慮した再学習

同じ分類（Fターム）でも、年代によって技術分野が変化してきており、教師文献と最新の特許文献との間に内容の乖離が発生するため、付与精度が低下していく可能性が高い。そのため、付与精度の低下を検知する仕組みを構築し、付与精度が一定以上低下した場合は、再学習等を検討する等の対応が必要である。

(4) 付与根拠箇所が図面に存在する場合への対応

本検証において、付与根拠箇所は特許文献明細書のテキスト部分についてのみ解析対象としたが、場合によってはテキスト以外の図面に根拠がある場合も想定される。さらなる精度向上化のためには、テキストに加え、

図面も利用できる機械学習モデルを検討する必要がある。

(5) 教師文献の少ないテーマ・分類における機械学習

今回対象とした12テーマは、比較的教師文献が存在しているテーマ及び分類を対象としているが、中には、テーマ及び分類に対する教師文献が少ないものもある。そのようなテーマ及び分類は、分類付与精度及び付与根拠箇所推定精度が低い可能性がある、そのため、精度向上においては、類似する技術分野の文献を教師データとして活用し、少ない教師データでも精度が向上することが可能な転位学習等を活用した方式を検討する必要がある。

3 おわりに

本調査事業の結果から、一定の精度で付与根拠箇所の推定が可能であるという知見を得た。一方で、実際の分類付与業務への適用には、学習データの収集や、学習方法の工夫など、多くの改善点についても知見を得た。引き続き、付与精度の向上など、今後の活用に向けた検討を進めてまいりたい。

参考文献

- 1 特許庁における人工知能技術の活用（平成28年度
の取組と今後のアクションプラン）平成29年4
月 特許庁
https://www.jpo.go.jp/torikumi/t_torikumi/pdf/ai_action_plan/01.pdf
- 2 富永泰規、外国特許文献への分類付与に関する機
械学習活用可能性調査について、Japio YEAR
BOOK 2017、2017、p. 212-216
- 3 Zichao Yang, et al. Hierarchical Attention
Networks for Document Classification, In
NAACL-HLT, 2016, pp. 480-1489
- 4 柳瀬利彦ほか、Neural Attention Model を用いた
観点付き評判分析、The 30th Annual Conference
of the Japanese Society for Artificial Intelligence、
2016.
- 5 Yuta Koreeda, et al. Neural Attention Model
for Classification of Sentences that Support
Promoting/ Suppressing Relationship. In
ArgMining, 2016, pp. 76-81.
- 6 榊原隆文ほか、付与根拠箇所推定に基づく特許文書
へのFターム付与、言語処理学会第23回年次大
会発表論文集、2017.
- 7 Thomas G. Dietterich, et al., Solving the
multiple instance problem with axis-parallel
rectangles, Artificial intelligence, 1997, 89
(1): 31-71.
- 8 Piotr Bojanowski, et al. Enriching Word Vectors
with Subword Information. In TACL 2017,
vol. 5(0), pp. 135-146.
- 9 Armand Joulin, et al. Bag of Tricks for Efficient
Text Classification. In EACL, 2017, pp. 427-
431.
- 10 Kazuma Hashimoto, et al., A Joint Many-Task
Model: Growing a Neural Network for Multiple
NLP Tasks. In EMNLP, 2017, pp. 446-456.

