Technology Reports
自然言語処理　機械翻訳　文書分類
豊かで多様性のある社会を支援するAI

機械翻訳における利用ドメインの自動推定

サービスイノベーション部　村上 聡一朗（むらかみ そういちろう）
澤山 熱気（さわやま あつき）　中村 俊允（なかむら としみつ）
松岡 保静（まつおか ほうせい）　内田 渉（うちだ わたる）

機械翻訳においては，飲食業や宿泊施設，交通機関といったさまざまな利用ドメインごとの会話の内容や言回し，対訳が異なるため，翻訳性能の向上のためには，そのドメインに適したコーパスを用いて，各ドメイン専用の機械翻訳エンジンを作成することが一般的である．しかし，複数のドメインを対象とした場合，ユーザが複数の翻訳エンジンを選ぶ必要があり，利便性を損なうおそれがある．そこでドコモでは，機械翻訳エンジンの利用ドメインをユーザが入力したテキストから自動推定する技術を開発した．これにより，入力されたテキストに対して最適な機械翻訳エンジンを自動選択することが可能となった．

1. まえがき

近年，日本を訪れる外国人旅行者の数は右肩上がりに増え続け，インバウンド需要が急増している．これを背景に，レストランなどの飲食業や宿泊施設，交通機関などでは，外国人旅行者との円滑なコミュニケーションを図るため，音声認識技術と機械翻訳技術を活用した音声翻訳サービスを導入している．さらに，旅行中の病気や怪我の際に利用される医療機関においても，診察・治療時の意思の疎通のために音声翻訳サービスが導入され，その利用シーンは広範にわたってきている．

そのような中，機械翻訳の分野では，ニューラル機械翻訳*1が注目されている［1］［2］．ニューラル機械翻訳とは，対訳コーパス*2を用いて大規模なニューラルネットワーク（NN：Neural Network）*3の学習を行うもので，従来の統計的機械翻訳よりも流暢で正確な翻訳が可能となった［3］．

©2019 NTT DOCOMO, INC.
本誌掲載記事の無断転載を禁じます．

＊1　ニューラル機械翻訳：機械学習の手法の1つであるニューラルネットワークを利用した機械翻訳技術．
＊2　コーパス：テキストや発話を大量に収集してデータベース化した言語資料．

ニューラル機械翻訳では，ドメイン*4に適した大量かつ高品質な対訳コーパスを用いることで，そのドメインにおける翻訳性能を向上させることができる．そのため，ニューラル機械翻訳をベースとした音声翻訳サービスでは，各ドメイン専用の機械翻訳エンジン*5を用意することが一般的である．しかし，インバウンド需要の急増に伴い，音声翻訳サービスの利用が想定されるドメインも増える．そして，それぞれのドメインに適した機械翻訳エンジンが必要となるが，ユーザにとって利用するドメインごとにどの機械翻訳エンジンを使うか選択するのは煩雑である．

そこでドコモでは，入力したテキストから，そのドメインを自動で識別するドメイン自動推定技術を開発した．これにより，ユーザが音声やキーボードを介して入力した文のドメインを推定し，そのドメインに適した機械翻訳エンジンを自動的に選択して翻訳を行うことができる．

本稿では，音声翻訳サービスの利用シーンを自動推定するドメイン推定技術について解説する．

2. 音声翻訳サービスの課題

海外旅行や外国人旅行者との接客に特化した音声翻訳サービスとして，国立研究開発法人情報通信研究機構（NICT：National Institute of Information and Communications Technology）の音声翻訳アプリVoiceTra®*6やログバー社の接客用オフライン翻訳デバイスili®*7，ソースネクスト社の翻訳デバイスPOCKETALK®*8などがある．また，ドコモでは，海外旅行時のコミュニケーションや日本国内での外国人旅行者とのコミュニケーションが円滑となるように，スマートフォン用翻訳アプリ「はなして翻訳」を提供している．

このように音声翻訳サービスは，多様な展開をみせ，機械翻訳が幅広いドメインで利用されるようになった．一方，飲食業や宿泊施設，交通機関などのドメインで翻訳が必要とされる内容，使われる単語や言回し，およびその対訳はそれぞれ異なるため，ドメインごとに適した専用の機械翻訳エンジンを導入し，その翻訳性能の向上を図ってきた．しかしながら，ユーザが音声翻訳サービスを利用する際に，サービスごとに専用の機械翻訳エンジンを選択するのは手間がかかる．そこで，ユーザが入力したテキストのドメインを推定し，最適な機械翻訳エンジンを自動で選択することができれば，それを回避することができると考えた．

3. ドメイン自動推定

今回，開発したドメイン自動推定技術では，ユーザが入力したテキストから特徴抽出を行い，機械学習*9による文書分類を行うことで，入力テキストの適切なドメインを推定する．

本システムの概要図を**図1**に示す．図1は，ユーザの入力文が，文書分類器によって，あらかじめ定められたドメインのいずれかに分類され，そのドメイン専用の翻訳エンジンへ翻訳リクエストを送信する流れを表している．ここでいう文書分類器とは，テキストをあらかじめ定められた分類先のいずれかに分類する装置のことである．

3.1 文書分類としてのドメイン自動推定

本技術では，入力テキストのドメインを推定するために，文書分類を行う．文書分類とは，音声翻訳サービスに入力されたテキストをあらかじめ定義したラベルのいずれかに分類することであり，ラベルは，音声翻訳サービスが対象とする飲食業や宿泊施設，交通機関などのドメインのことを指す．一般的に自然言語処理の分野では，文書とラベルがペアになった学習データを用いて機械学習モデルを訓練することで，文書分類器*10を構築する．

---

*3 ニューラルネットワーク（NN）：人間の脳内にある神経細胞（ニューロン）とそのつながりを数式的なモデルで表現したもの．入力層，出力層，中間層から構成され，ニューロンや層の数，層の間のつながりの強さを変更することにより複雑な関数近似を行うことができる．

*4 ドメイン：機械翻訳の利用シーンに相当するもの．

*5 機械翻訳エンジン：機械翻訳を行うソフトウェアのこと．

*6 VoiceTra®：（国研）情報通信研究機構の登録商標．

*7 ili®：㈱ログバーの登録商標．

*8 POCKETALK®：ソースネクスト㈱の登録商標．

文書分類器の分類例を**図2**に示す．この例では，音声翻訳サービスに入力された「再診の受付は5番カウンターになります」というテキストを文書分類器で特徴抽出し，あらかじめ定義したドメインのラベルの中から機械学習手法を用いて適切なドメイン「医療機関」を推定している．

3.2　文書分類器の学習データ

機械学習手法を用いた文書分類器を訓練するために，音声翻訳サービスの入力テキストとドメインのラベルがペアになった学習データを使用する．

一般的に機械学習手法では，学習データが多ければ多いほどモデル性能が向上する．また，学習データを構築する際には，ドメインによってペアとなるデータが多い・少ないなどの不均衡な学習データにならないように注意が必要である．これは，データが少ないドメインのテキストの分類精度が低下するという過学習の問題を防ぐためである．

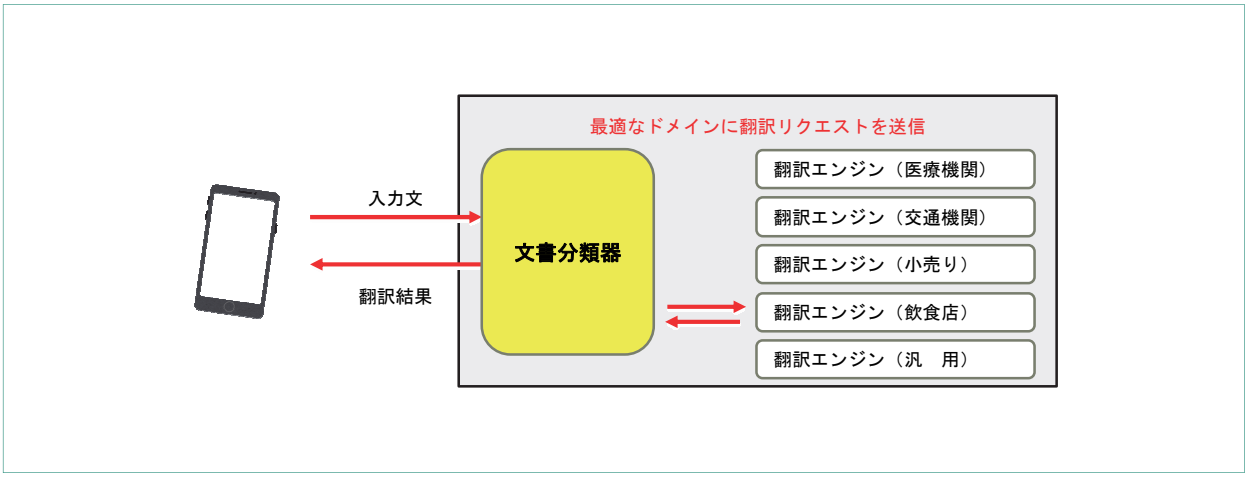

図1　システム概要図

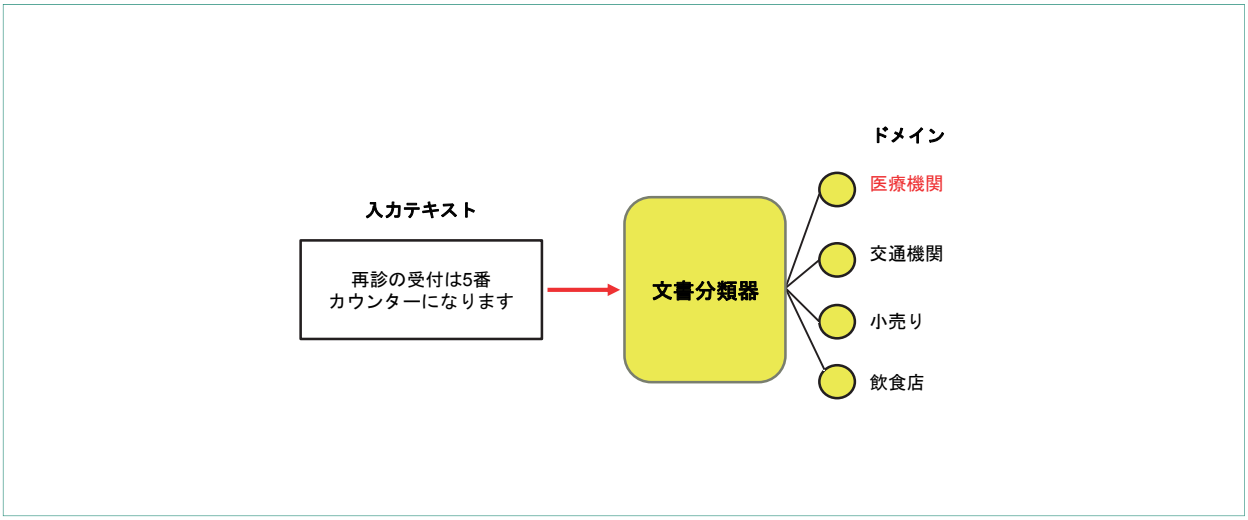

図2　文書分類器の概要図

＊9　機械学習：人間が，知覚，経験から知識や判断基準，動作などを獲得していくように，コンピュータにデータから知識や判断基準，動作などを獲得させる技術．

＊10　分類器：入力を，その特徴量に基づいてあらかじめ定められた分類先のいずれかに分類する装置．

文書分類器の学習データの例を**表1**に示す．この例では，表中のテキスト「めまいがする場合に，冷や汗をかいたり悪寒を感じたりしますか？」のドメインとして「医療機関」がラベル付けされていることを表している．

3.3　文書分類器の機械学習手法

本システムで使用されるLSTM（Long-Short Term Memory）[4]を用いた文書分類器について解説する．LSTMとは，NNに再帰結合*11を導入したRNN（Recurrent NN）の一種であり，可変長のテキストを取り扱う自然言語処理の分野で広く使われている．順伝播型NN*12と再帰型NNの概要図を**図3**に示す．ここで，LSTMおよびRNNは，再帰型NNに相当する．

LSTMやRNNのような再帰型NNでは，NNの中間層において，1つ前の時刻t-1の内部状態ベクトルを次の時刻tで引き継ぐことができるため，テキス

表1　文書分類器の学習データの例

テキスト	ドメイン
先生に診てもらえますか？	医療機関
めまいがする場合に，冷や汗をかいたり悪寒を感じたりしますか？	医療機関
ICカードのチャージはできませんので事前にお願い致します．	交通機関
天候や道路状況によっては時刻通りに到着しない場合もあります．	交通機関
返品・交換の際はレシートか領収書をお持ちください．	小売り
歯磨き粉はどこにありますか？	小売り
日本酒，焼酎を除いたメニューを全て飲み放題でお楽しみいただけます．	飲食店
ジュースは全て100%で砂糖などは使用しておりません．	飲食店

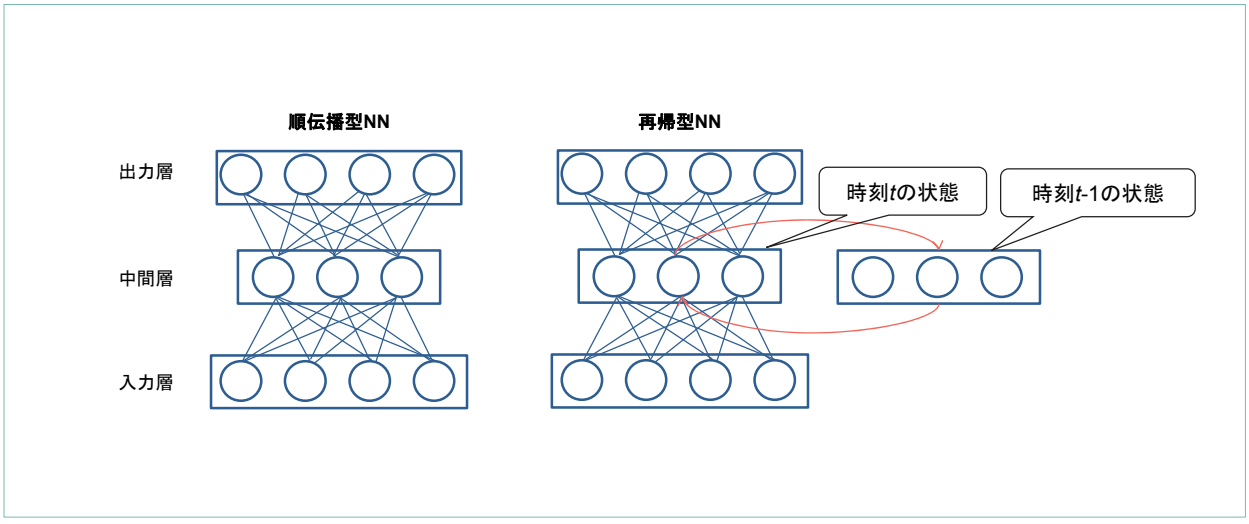

図3　順伝播型NNと再帰型NN

*11 再帰結合：再帰的に結合されること．
*12 順伝播型NN：ネットワークに再帰する結合を持たず，入力層，中間層，最終層の順に，単一方向へのみ信号を伝搬するニューラルネットワークのこと．

トのような可変長の入力を柔軟に取り扱うことが可能となる．また，テキストを時系列に入力できるため，テキストの文脈情報を文脈ベクトルとして，固定長のベクトルで表現することができる．つまり，再帰型NNを用いることで，入力テキストから文脈を表す特徴量*13を抽出することが可能となる．

LSTMを用いた文書分類器の判定例を**図4**に示す．ここで，LSTMのベクトルの次元数は200とする．LSTMを用いた文書分類器では，NNを用いて入力テキストから作成した固定長の文脈ベクトルを基に，入力テキストがどのドメインに最も適合するかを判定する．まず，「頭が痛いです」という入力文を形態素解析*14により単語分割した入力単語列「頭　が　痛い　です　<EOS>」を獲得する．ここで，「<EOS>」は文の終わりを表す疑似トークン*15である．次に，入力単語列の各単語をベクトル化した200次元の単語ベクトルをLSTMに1つずつ入力し，入力文の文脈情報を表す文脈ベクトルを計算する．最後に，この文脈ベクトルを基にSoftmax関数*16を用いて，入力文がどのドメインに最も適合するかを表す尤度を算出し，ドメイン推定を行う．

3.4　ドメイン推定精度

テキストとラベルからなるペアデータを用いて，LSTMを用いた文書分類器の訓練を行い，テストデータに対する分類精度を計測した．実験では，ドメインとして医療機関，交通機関，小売り，飲食店を定義し，各ドメインのテストデータを1,000文ずつ用意した．

ドメイン推定精度を**表2**に示す．各ドメインの分類精度（*F*値*17）をみると，おおむね高精度に分類できていることが分かった．また，1文あたりのドメイン推定にかかる平均処理時間は，約12msであり，実用に問題ない程度の処理時間でドメイン推定を行うことができることが分かった．

3.5　利用例

次に，本システムがドメイン推定技術を用いて機械翻訳を行う利用例について解説する（図1）．入力テキストとして，

「このレストランではカリフォルニア産の高級ワインが召し上がれます．」

を用いる．文書分類器を用いて，この文書のドメインの自動推定を行うと，「飲食店」というドメインに

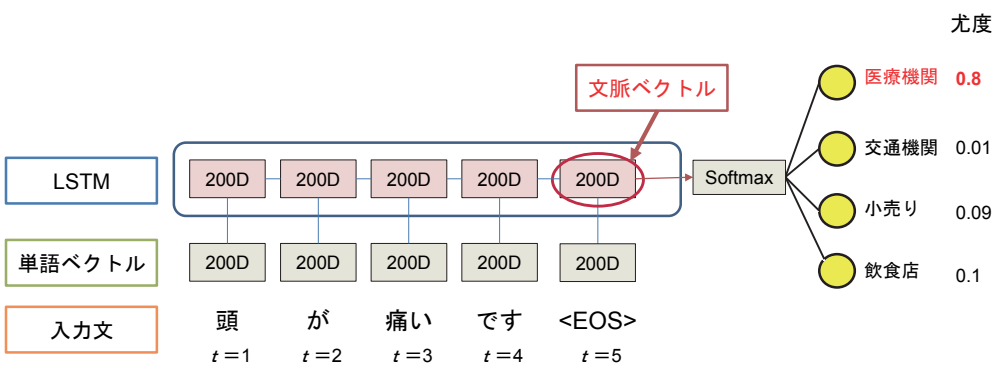

図4　LSTMを用いた文書分類器

*13 特徴量：データから抽出される，そのデータを特徴づける値のこと．
*14 形態素解析：自然言語で書かれたテキストを言語で意味を持つ最小単位である形態素に分割し，それぞれの品詞などを判別する操作．
*15 トークン：テキストの最小単位として扱われる文字や文字列のこと．
*16 Softmax関数：NNの出力の合計を1.0となるよう正規化し，確率値を計算する際に用いられる．
*17 *F*値：正確性と網羅性の総合的な評価の際に利用される尺度であり，適合率と再現率の調和平均によって算出される．

表2　ドメイン推定精度

ドメイン	LSTM			事例数
	適合率	再現率	*F*値	
医療機関	0.99	0.92	0.95	1,000
交通機関	0.95	0.95	0.95	1,000
小売り	0.87	0.94	0.91	1,000
飲食店	0.92	0.92	0.92	1,000

※平均処理時間：1文当り12ms

推定される．次に，この入力テキストを飲食店ドメイン用の機械翻訳エンジンで翻訳を行うと以下のように翻訳される．

“You can enjoy California high-quality wine at this restaurant.”

また，この文を汎用の機械翻訳エンジンで翻訳した場合，以下のようになる．

“This restaurant has a high quality wine in California.”

飲食店ドメイン専用の機械翻訳エンジンと汎用の機械翻訳エンジンの翻訳結果を比較すると，専用の機械翻訳エンジンの方がそのドメインで通常使用されるであろうフレーズを用いて流暢に翻訳できることが分かる．

このように，ドメイン自動推定技術を用いて，入力テキストのドメインに対して最適な機械翻訳エンジンで翻訳をすることで，より高品質な翻訳を行うことが可能である．

4. あとがき

本稿では，さまざまなドメインで利用されている音声翻訳サービスにおいて，ユーザが利用するドメインに合った機械翻訳エンジンで翻訳を行うために，ドメイン自動推定を用いて最適な機械翻訳エンジンを自動的に選択する技術を解説した．

今後の展望として，より高精度なドメイン推定技術の開発やテキスト以外の情報を用いたドメイン推定技術の開発などが考えられる．

文　献

[1] I. Sutskever, O. Vinyals and Q. V. Le: “Sequence to Sequence Learning with Neural Networks,” Advances in neural information processing systems, pp.3104-3112, 2014.

[2] D. Bahdanau, K. Cho and Y. Bengio: “Neural Machine Translation by Jointly Learning to Align and Translate,” In Proc. of the 3rd International Conference on Learning Representations, 2014.

[3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N Gomez, L. Kaiser and I. Polosukhin: “Attention is All You Need,” In Proc. of Advances in Neural Information Processing Systems 30, pp.5998-6008, 2017.

[4] S. Hochreiter and J. Schmidhuber: “Long Short-Term Memory,” Neural computation, Vol.9, No.8, pp.1735-1780, 1997.