

AI 特許紹介(16)
AI 特許を学ぶ！究める！
～Sequence to Sequence Learning～

2020年6月10日
河野特許事務所
所長 弁理士 河野英仁

「AI 特許紹介」シリーズは、注目すべき AI 特許のポイントを紹介します。熾烈な競争となっている第4次産業革命下では AI 技術がキーとなり、この AI 技術・ソリューションを特許として適切に権利化しておくことが重要であることは言うまでもありません。

AI 技術は Google, Microsoft, Amazon を始めとした IT プラットフォーマ、研究機関及び大学から毎週のように新たな手法が提案されており、また AI 技術を活用した新たなソリューションも次々とリリースされています。

本稿では米国先進 IT 企業を中心に、これらの企業から出願された AI 特許に記載された AI テクノロジー・ソリューションのポイントをわかりやすく解説致します。

1.概要

特許権者 Google

出願日 2015年6月4日

登録日 2019年1月15日

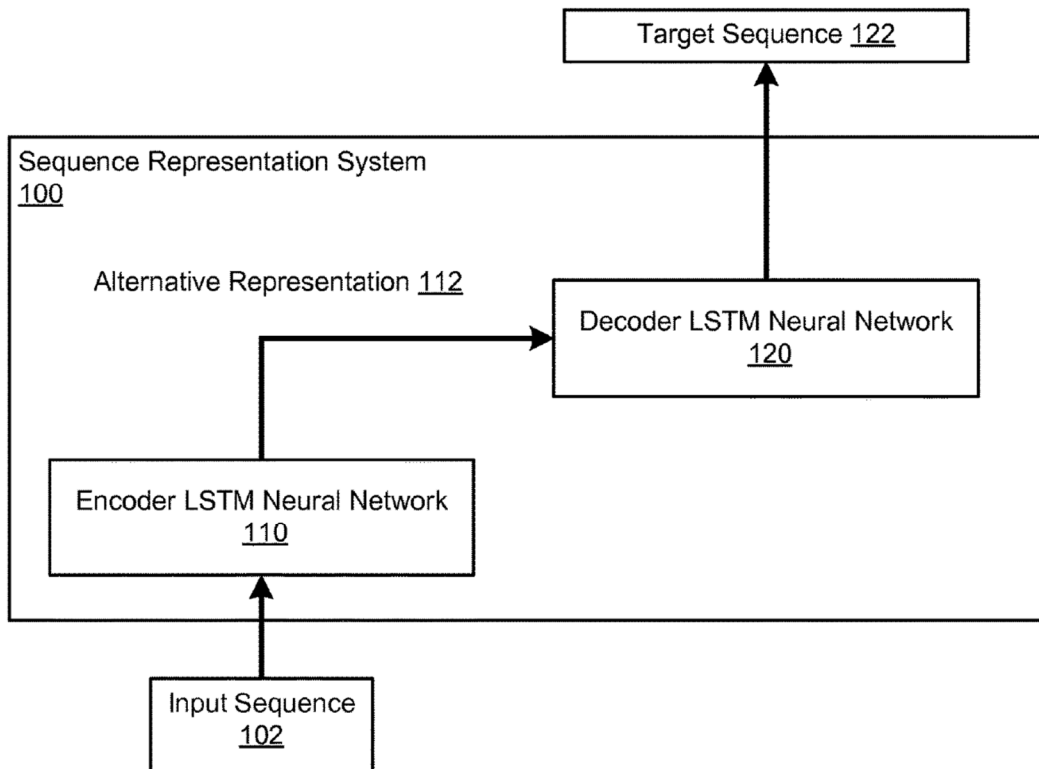
登録番号 US10181098

発明の名称 ニューラルネットワークを使用した入力シーケンスの表現の生成

098 特許は、入力シーケンスから LSTM ニューラルネットワークを使用して出力シーケンスを生成するアイデアに関する。

2.特許内容の説明

下記図は、シーケンス表現システム 100 の構成を示す説明図である。



シーケンス表現システム 100 は、入力シーケンス 102 を受け取り、入力シーケンス 102 をターゲットシーケンス 122 に変換する。各ターゲットシーケンス 122 は、入力シーケンス 102 の表現、たとえば、異なる形式の入力シーケンスの表現である。

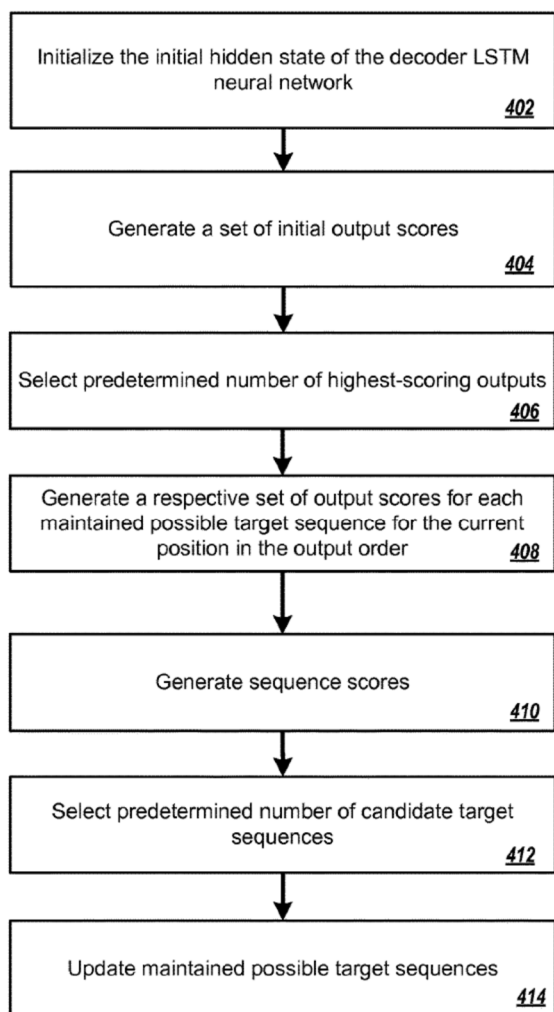
例えば、入力シーケンス 102 が原文の単語のシーケンス、例えば、文または句である場合、シーケンス表現システム 100 によって生成されるターゲットシーケンス 122 は、入力シーケンスのターゲット言語への翻訳、すなわち、元の言語の単語のシーケンスを表すターゲット言語の単語のシーケンスである。

別の例として、入力シーケンス 102 が書記素シーケンス、例えば、シーケンス{g, o, o, g, l, e}である場合、シーケンス表現システム 100 によって生成されたターゲットシーケンス 122 は、入力シーケンスの音素の代替表現、例えば、シーケンス{g, uh, g, ax, l}である。

一般に、シーケンス表現システム 100 によって受信された入力シーケンスおよびシーケンス表現システム 100 によって生成されたターゲットシーケンスは、可変長シーケンス、すなわち、さまざまな数の入力および出力をそれぞれ含むことができるシーケンスである。

シーケンス表現システム 100 は、エンコーダ長期短期メモリ (LSTM) ニューラルネットワーク 110 およびデコーダ LSTM ニューラルネットワーク 120 を含む。シーケンス表現システム 100 は、エンコーダ LSTM ニューラルネットワーク 110 を使用して入力シーケンスを処理して、入力シーケンス 102 の代替表現を生成する。

以降の処理を下記フローチャートを用いて説明する。



システムは、入力シーケンスの代替表現に対して、デコーダ LSTM ニューラルネットワーク 120 の初期の隠れ状態を初期化する (ステップ 402)。システムは、初期の隠れ状態に従って、デコーダ LSTM ニューラルネットワーク 120 を使用して初期出力スコアのセットを生成する (ステップ 404)。

システムは、初期スコアに従って、所定の数の最高スコアの可能な出力を選択する (ステップ 406)。システムは、選択された可能な出力ごとにそれぞれの可能なターゲットシーケンスを生成する。システムは、出力順序における現在の位置について維持された

可能なターゲットシーケンスごとに、それぞれの出力スコアのセットを生成する（ステップ 408）。

システムは、維持された可能性のあるターゲットシーケンスごとに、可能性のあるターゲットシーケンスから生成できる可能性のある各ターゲットシーケンスのそれぞれのシーケンススコアを生成する（ステップ 410）。システムは、最高のシーケンススコアを有する所定数の候補ターゲットシーケンスを選択する（ステップ 412）。

システムは、維持された可能なターゲットシーケンスを選択された候補ターゲットシーケンスになるように更新する（ステップ 414）。選択された候補ターゲットシーケンスが文末(end-of-sentence)トークンで終了すると、システムは選択された候補ターゲットシーケンスをビームから削除する。つまり、システムは、選択された候補ターゲットシーケンスへの追加出力の追加を停止し、文末トークンが追加される前に選択された候補ターゲットシーケンスを、最終的に可能なターゲットシーケンスと見なす。

システムは、維持される可能性のある各ターゲットシーケンスが確定されるまで、プロセス 400 のステップ 408～414 を繰り返す。

3.クレーム

098 特許のクレーム 1 は以下の通りである。

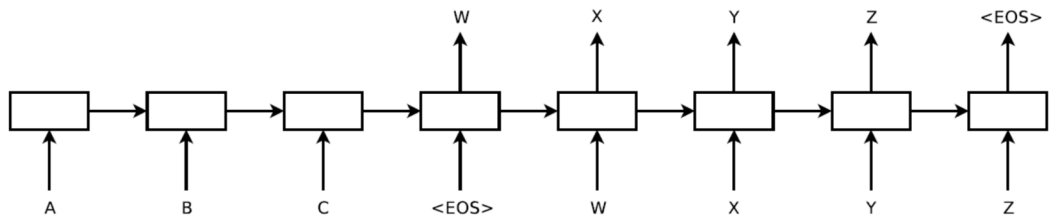
- 1 つまたは複数のコンピュータによって実行される方法において、
 - 入力順序に従って配置された複数の入力を含む入力シーケンスを取得し、
 - 入力シーケンスを入力シーケンスの代替表現に変換するために、第 1 長期短期記憶 (LSTM) ニューラルネットワークを使用して入力シーケンスを処理し、
 - 入力シーケンスのターゲットシーケンスを生成するために第 2LSTNM ニューラルネットワークを使用して入力シーケンスの代替表現を処理し、前記ターゲットシーケンスは、出力順序に従って配置された複数の出力を含む。

4. Seq2Seq に関する論文

ニューラルネットワークを用いた Seq2Seq 学習に関する論文¹が Google の Ilya Sutskever 氏らにより発表されている。

論文には Seq2Seq のモデルが紹介されている。

¹ Ilya Sutskever, Oriol Vinyals, Quoc V. Le “Sequence to Sequence Learning with Neural Networks” arXiv:1409.3215v3 [cs.CL] 14 Dec 2014



本モデルは、入力シーケンス「ABC」を読み取り、出力シーケンスとして「WXYZ」を生成する。モデルは、EOS(end-of-sentence)を出力した後、予測を停止する。具体的には、LSTMは「A」、「B」、「C」、「<EOS>」の表現を計算し、この表現を使用して「W」、「X」、「Y」、「Z」、「<EOS>」の確率を計算する。

上記図は概念図であり、実際に構築したモデルは以下の3つの特徴を有する。

1つ目には、098 特許で示された2つの異なる LSTM を使用していることである。1つは入力シーケンス用のエンコーダであり、もう1つは出力シーケンス用のデコーダである。これにより、無視できるほどの計算コストで、モデルパラメーターの数を増加させることができ、LSTM を複数の言語ペアで同時にトレーニングすることが可能となる。

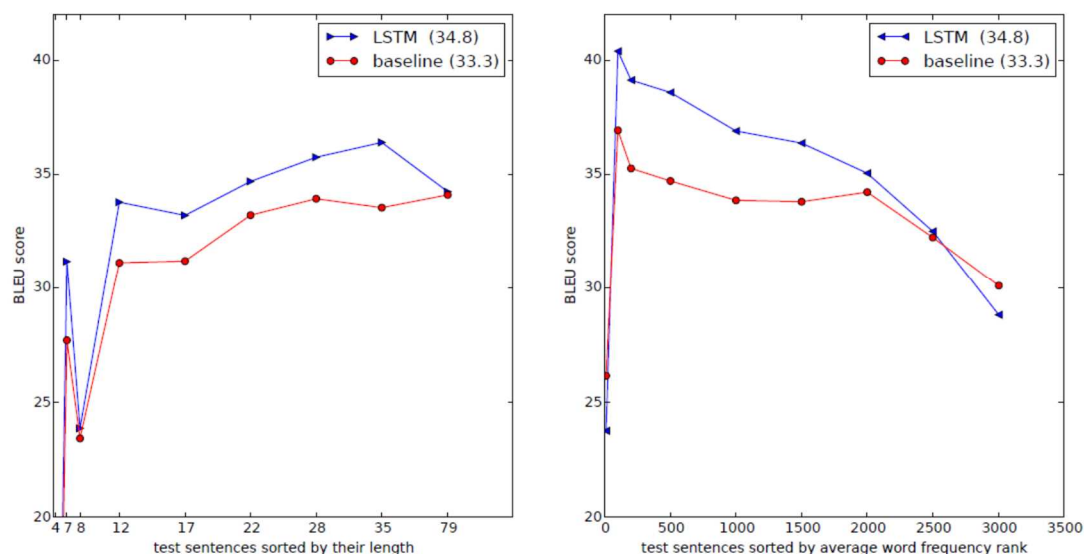
第2に、深層 LSTM として、4層の LSTM を選択している。

第3に、入力文の単語の順序を逆にすることに、非常に価値があることを見出している。

例えば、文章 a, b, c から、その翻訳文である α , β , γ をマッピングする代わりに、LSTM は、c, b, a から α , β , γ をマッピングする。

このように、a は α に近接し、b は β にかなり近接する。これにより、SGD(Stochastic Gradient Descent 確率的勾配降下法)にとって、入力と出力の間の「通信を確立」することが容易となる。このシンプルなデータ変換により、LSTM の性能を大幅に向上させることができる。

論文には試験結果が示されている。



左のグラフは、システムのパフォーマンスを文の長さの関数として示したものであり、赤色がベースライン、青色が LSTM である。x 軸は、長さで並べ替えられたテスト文に対応し、実際のシーケンスの長さでマークされており、縦軸は機械翻訳の評価に用いられる BLEU スコアである。35 語未満の文の劣化はなく、最長の文のわずかな劣化のみが認められる。

右のグラフは、次第に珍しい単語を含む文に対する LSTM のパフォーマンスを示している。x 軸は、「平均単語頻度ランク」でソートされたテスト文に対応している。

以上

著者紹介

河野英仁

河野特許事務所、所長弁理士。立命館大学情報システム学博士前期課程修了、米国フランクリンピアースローセンター知的財産権法修士修了、中国清華大学法学院知的財産夏季セミナー修了、MIT(マサチューセッツ工科大学)コンピュータ科学・AI 研究所 AI コース修了。

[AI 特許コンサルティング](#)、[医療 AI 特許コンサルティング](#)の他、米国・中国特許の権利化・侵害訴訟を専門としている。著書に「世界のソフトウェア特許(共著)」、「FinTech 特許入門」、「[AI/IoT 特許入門 2.0](#)」がある。