

AI 特許紹介(12)

～トピックベースのシーケンスニューラルネットワーク～

2020年3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 2016年8月26日

登録日 2018年9月25日

登録番号 US10083169

発明の名称 トピックベースのシーケンスモデリングニューラルネットワーク

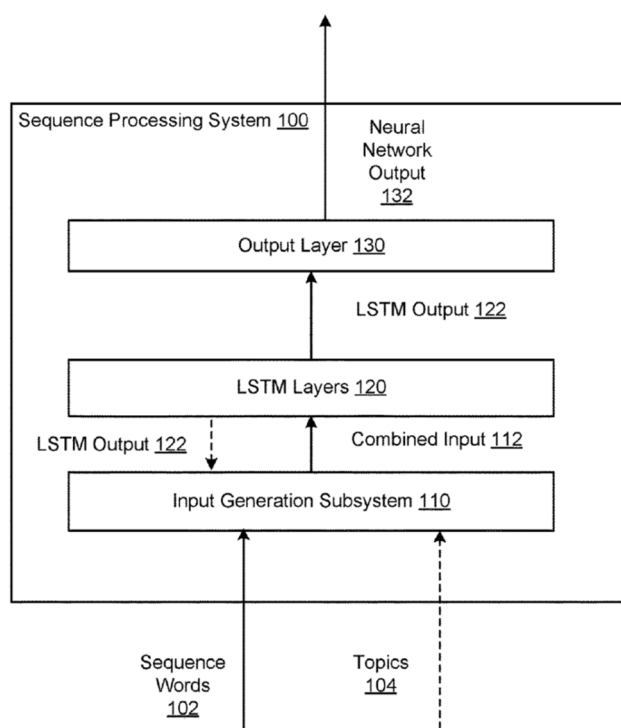
169 特許は、文章中の単語に加えて単語のトピックを入力要素の一つに加え、次に出現する単語、文章またはトピックを予測するニューラルネットワークに関する発明である。

2.特許内容の説明

リカレントニューラルネットワークは、入力シーケンスを受け取り、入力シーケンスから出力シーケンスを生成するニューラルネットワークである。特に、リカレントニューラルネットワークは、現在のタイムステップで出力を計算する際に、前のタイムステップのネットワークの内部状態の一部または全てを使用する。

リカレントニューラルネットワークの一つとして、1 つまたは複数の LSTM メモリブロックを含む長期 (LSTM) ニューラルネットワークがある。各 LSTM メモリブロックには、入力ゲート、忘却ゲート、および出力ゲートを含む 1 つ以上のセルが含まれている。

169 特許では、単語シーケンス内の各単語について、単語のトピックベクトルを決定し、単語とトピックベクトルに基づき結合入力を生成する。そして、生成した結合入力に対応するニューラルネットワーク出力を生成する。出力としては次に出現する単語、文章、トピックが含まれる。



シーケンス処理システム 100 は、入力生成サブシステム 110、ロングショートタームメモリ (LSTM) ニューラルネットワーク層 120、および出力層 130 を含む。

単語シーケンス 102 内の各単語について、入力生成サブシステム 110 は、単語および単語のトピックベクトルから結合入力を生成する。単語のトピックベクトルは、その単語が属するテキストセグメントのトピックの数値表現である。たとえば、猫を説明するテキストセグメントは、そのトピックベクトルとして、単語「猫」をマッピングするベクトルを有する。

入力生成サブシステム 110 は、単語シーケンス 102 内のすべての単語に対して同じトピックベクトルを使用することができ、また単語シーケンス 102 が複数の文を含む場合、単語シーケンス 102 内の同じ文のすべての単語に対してのみ同じトピックベクトルを使用することもできる。

単語シーケンス 102 内の各単語について、LSTM層 120 は、その単語について入力生成サブシステム 110 によって生成された結合入力を受け取り、LSTM層 120 の現在の隠れ状態に従って結合入力を処理して生成する。

出力層 130 は LSTM 層 120 出力を処理し、次の単語、トピックまたは文章を予測する。

3.クレーム

169 特許のクレーム 1 は以下の通りである。

1. 第 1 の順序に従って配置された第 1 の単語シーケンスを受信し、該第 1 の単語シーケンスは複数の単語文を含み、複数の文は入力順序であり、

第 1 の単語シーケンスにおける各単語に対し、第 1 の順序で最初の単語から開始し、単語に関連付けられているトピックベクトルを決定し、

単語とトピックベクトルから結合入力を生成し、

単語のそれぞれのシーケンスモデリング出力を生成するために、1 つ以上のシーケンスモデリングレイヤーを通じて結合入力を処理し、

第 1 の単語シーケンスのニューラルネットワーク出力を生成するために、出力層を介してそれぞれのシーケンスモデリング出力の 1 つ以上を処理し、

入力順序の最初の文の後の各文の各単語について、単語に関連付けられているトピックベクトルの決定は、少なくとも部分的には、入力順序の文の直前の文の最後の単語のシーケンスモデリング出力に基づく

方法。

4. CLSTM の論文

CLSTM に関しては、Shalini Ghosh らが 2016 年 5 月に論文¹を発表している。この論文は、CLSTM (Contextual LSTM)を紹介しており、トピックなどの文脈を LSTM に組み合わせることで次に出現する単語、文章、トピックを従来の LSTM より高精度で予測する技術及び実験結果について以下の通り論じている。

¹ Shalini Ghosh “Contextual LSTM (CLSTM) models for Large scale NLP tasks”

LSTMは、時系列データを処理するリカレントニューラルネットワークの一種である。LSTMモデルには複数のLSTMセルが設けられ、各LSTMセルはニューラルネットワークのデジタルメモリをモデル化している。

LSTMモデルには、LSTMが時間をかけて情報を保存およびアクセスできるようにするゲートが設けられている。たとえば、入出力ゲートはセルの入出力を制御し、忘却ゲートはセルの状態を制御する。

LSTMモデルは、入力から1ホットエンコードでエンコードされた単語を取得し、それらを埋め込みベクトルに変換し、単語ベクトルを1つずつ消費する。モデルは、すでに見られた一連の単語を前提として、次の単語を予測するようにBPTT(Backpropagation Through Time)等を用いてトレーニングされる。

LSTMセルのオペレーションを表す方程式に、トピックベクトル \mathbf{T} を入力ゲート、忘却ゲート、セル、および出力ゲートに追加する。

$$\begin{aligned}
 i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i + \mathbf{W}_{Ti}\mathbf{T}) \\
 f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f + \mathbf{W}_{Ti}\mathbf{T}) \\
 c_t &= f_t c_{t-1} + i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c + \mathbf{W}_{Ti}\mathbf{T}) \\
 o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o + \mathbf{W}_{Ti}\mathbf{T}) \\
 h_t &= o_t \tanh(c_t)
 \end{aligned} \tag{1}$$

各式の太字が修正を加えた部分である。ここで、 i 、 f 、および o はそれぞれ入力ゲート、忘却ゲート、出力ゲート、 x は入力、 b はバイアス項、 c はセルメモリ、 h は出力である。

例として、入力ゲート方程式を考える。

$$\begin{aligned}
 i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \\
 &= \sigma([W_{xi} \ W_{hi} \ W_{ci} \ 1][x_t \ h_{t-1} \ c_{t-1} \ b_i]^T)
 \end{aligned} \tag{2}$$

トピック信号 \mathbf{T} を入力ゲートに追加すると、方程式は次のように変更される。

$$\begin{aligned}
 i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i + W_{Ti}T) \\
 &= \sigma([W_{xi} \ W_{Ti} \ W_{hi} \ W_{ci} \ 1][x_t \ T \ h_{t-1} \ c_{t-1} \ b_i]^T)
 \end{aligned} \tag{3}$$

式2及び3を比較すると、トピックベクトル \mathbf{T} をCLSTMセルに追加することは、単語

埋め込みベクトルとトピック埋め込みベクトルを連結するLSTMセルへの複合入力 $[x_i T]$ を考慮することと同等であることがわかる。

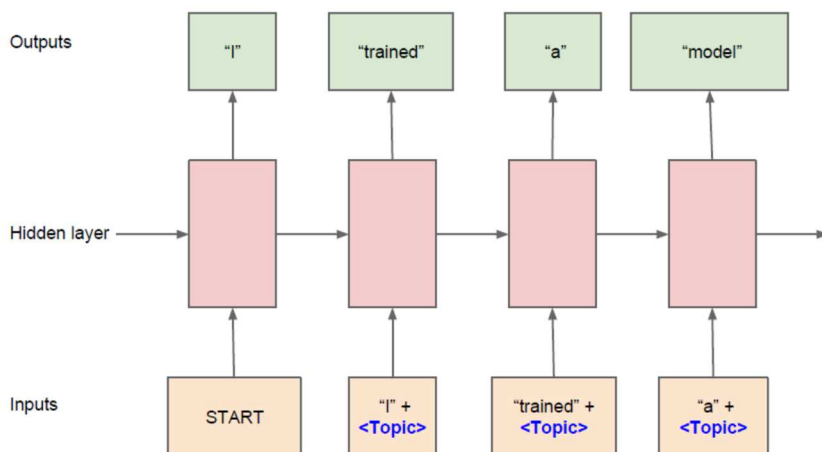
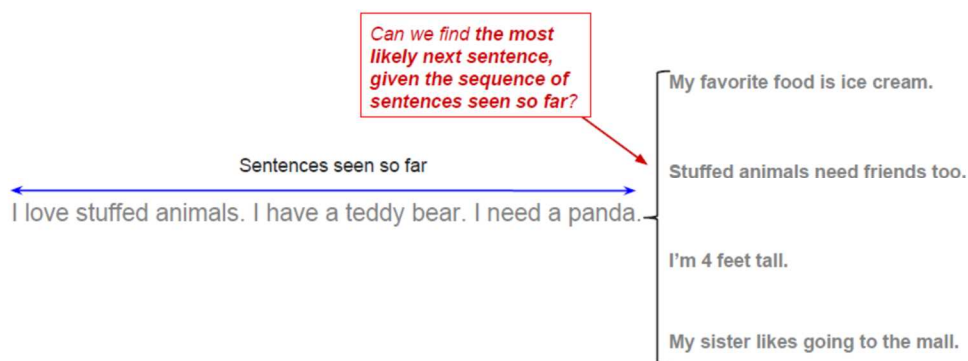


Figure 1: CLSTM model (<Topic> = topic input)

図1は、単語とトピックの両方の入力ベクトルを考慮したCLSTMモデルの概略図である。各入力セルには単語に加えてトピックが入力されている。トピックはセルごとに共通のトピックを有する場合もあれば、それぞれ異なるトピックを有する場合もある。



この CLSTM モデルを用いて、次に発生する単語、文章、またはトピックを予測する。上の例の場合、上から 2 番目の文章が予測出力される。

LSTM	CLSTM	Accuracy Increase
52% ± 2%	63% ± 2%	21% ± 9%

試験結果によれば、ベースライン LSTM モデルの平均精度は 52%、単語+文レベルのトピック機能を使用した CLSTM モデルの平均精度は 63%となり、CLSTM モデルは、LSTM モデルよりも精度が平均で 21%改善された。

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](#)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。

以上