

芝田 和聖†

静岡理科大学†

幸谷 智紀†

静岡理科大学†

1. 導入

TensorFlow(2)は、Google が開発した、ニューラルネットワーク用ライブラリである。sequence-to-sequence モデルを用いた、英語からフランス語への機械翻訳サンプルスクリプトがあり、RNN を用いた機械翻訳を行う事が出来る。

本稿では、このスクリプトを用いた日本語から英語への機械翻訳を試み、ハードウェア、パラメータ、RNN の活性化関数を変更し、学習速度や精度がどのように変化するか確認する。最後に、実用性向上へ向けた課題を挙げる。

2. 使用したソフトウェア、日英コーパス、翻訳指標

2.1 TensorFlow

TensorFlow とは、数値計算の為にライブラリである。Google がオープンソース化したライブラリで、2017年2月に正式版の TensorFlow1.0 がリリースされた。

本研究では、TensorFlow1.0 と TensorFlow1.2 の Python ライブラリを用いた。

2.2 sequence-to-sequence モデル

sequence-to-sequence モデルとは、入出力シーケンスの変換を行う、RNN を用いたモデルである。(5)の研究で考案された。多層 LSTM で、入力シーケンスをベクトルに変換し、もう一つの多層 LSTM で、ベクトルを出力シーケンスに変換する。

TensorFlow のチュートリアルで、Python で実装した sequence-to-sequence モデルのスクリプトが公開されている。スクリプトでは、LSTM(8)の他に GRU を用いる事が出来る。GRU は、(6)の研究で提案された RNN ユニットの。パラメータに、層ごとの RNN ユニットの size と、その層数の num_layers を指定する事が出来る。

本研究では、このスクリプトを用いて実験を行った。また、全ての実験において GRU を用いて実験を行った。

2.3 対訳コーパス

sequence-to-sequence モデルの学習には、1文ごとに対応関係を持つ対訳コーパスが必要である。本研究では、対訳コーパスに、日英法令対訳コーパス(3)と、ASPEC(1)を用いた。

2.4 機械翻訳の指標

本研究では、翻訳の指標に BLEU(7)を用いて、精度を評価した。指標の計算には自作のアプリケーションを用いた。

3. 実験と結果

3.1 ハード環境による学習速度のベンチマーク

ハードウェア環境を変更した際の学習速度の変化を、ベンチマークテストした。

Machine translation from Japanese to English using TensorFlow

† Shizuoka Institute of Science and Technology

この結果、GPU 非搭載のノート PC(Core i7-4600)は 20,042 秒、GeForce GTX1080 を搭載したデスクトップ PC(Core i7-6700K)は 3,150 秒を要した。結果から、ハードウェアは学習速度に大きく影響するといえる。

3.2 パラメータ変更による翻訳精度テスト

パラメータを変更した際に生じる精度の変化を、BLEU によって確認した。「size=256 num_layers=2」、「size=512 num_layers=3」、「size=256 num_layers=3」、「size=512 num_layers=3」のパラメータで、日英法令対訳コーパスを用いて 200,000step 学習を行った。これらのモデルで、コーパスに含まれる形態素数 11~20 と形態素数 21~40 の日本語文を翻訳した。この結果を、コーパスの英文を参照訳として BLEU を計算し、形態素数の分類別の平均と総平均を集計した。Fig.1 に、パラメータごとの BLEU のグラフを示す。

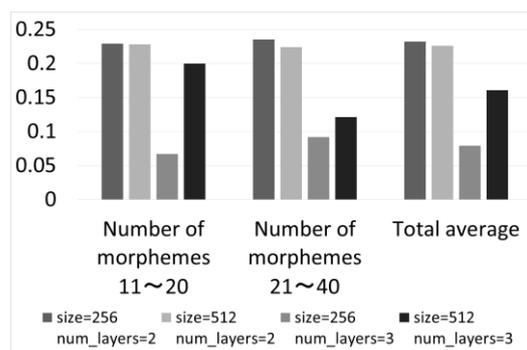


Fig.1 BLEU of each parameter

その結果、総平均が最も高かったのは、「size=256 num_layers=2」で、総平均は約 0.2324 だった。また、num_layers=3 のモデルは、num_layers=2 のモデルに比べ、全ての分類において、BLEU が低かった。

この結果から、モデルの層数を闇雲に増やしても、精度は上がらないと結論出来る。本研究で扱ったパラメータでは、num_layers=2 が最も適切なパラメータである。

3.3 学習ステップ数増加による翻訳精度テスト

学習ステップ数を増加させた際に生じる精度の変化を、BLEU によって確認した。「size=256 num_layers=3」のパラメータで、日英法令対訳コーパスを用いて 400,000step 学習を行った。この 200,000step 時点のモデルと、400,000step 学習したモデルで、3.2 の実験と同条件で翻訳し、BLEU を計算し集計した。Fig.2 に、学習ステップごとの BLEU のグラフを示す。

この結果、総平均で 400,000step の方が 200,000step よりも BLEU が高かったが、その差は小さかった。形態素数 11~20 においては、200,000step の方が 400,000step よりも BLEU が高かったが、形態素数 21~40 においては、400,000step の方が 200,000step よりも BLEU が高かった。

この結果は、ある程度学習したモデルの学習ステップ数を増やしても、精度向上に大きな効果は期待出来ない事を

示した。本研究では、二倍の学習ステップ数の学習を行ったが、総平均で大きな差は無く、二倍の時間をかけただけの効果は無かったと結論出来る。

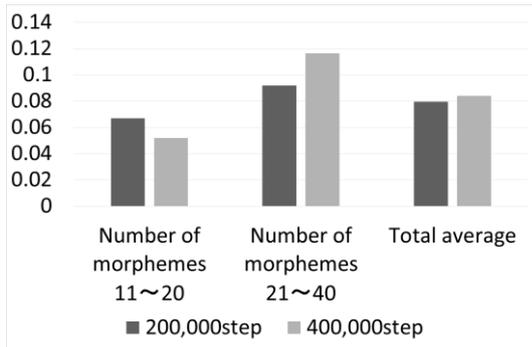


Fig.2 BLEU of each step

3.4 RNN の活性化関数変更による翻訳精度テスト

RNNの活性化関数を変更した際に生じる精度の変化を、BLEUによって確認した。ここまでは、活性化関数に tanh を用いていた。tanh とは、活性化関数の一般的な形状である、シグモイド曲線を描く関数のひとつである(9)。そこで、活性化関数に ReLU を用いる事で、精度がどのように変化するか検証した。(4)の実験で、ReLU を用いたニューラルネットワークは、tanh よりも学習が高速であった。

ここでは、「size=512、num_layers=2」と「size=512 num_layers=3」のパラメータで、tanh と ReLU を用いて、日英法令対訳コーパスを用いて 200,000step 学習を行った。これらのモデルで、3.2の実験と同条件で翻訳し、BLEU を計算し集計した。Fig.3 に、パラメータと活性化関数ごとの BLEU のグラフを示す。

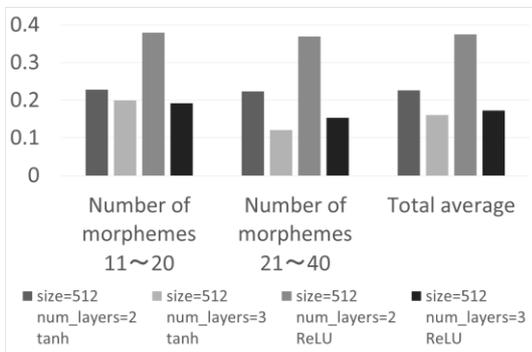


Fig.3 BLEU of each parameter and activation function

この結果、両パラメータで、ReLU を用いた方が、tanh を用いるよりも、BLEU が高かった。特に、「size=512 num_layers=2」で ReLU を用いると、総平均は 0.3749 となった。これは、本研究で最も高い BLEU であった。

この結果から、ReLU を用いる事で、精度向上が見込めると結論出来る。中でも特に、「size=512 num_layers=2」のパラメータのモデルの成績が良かった。これに対する参考記録として、sequence-to-sequence モデルによる英語からフランス語への翻訳(5)における最高の BLEU は 0.3481 で、条件は異なるが高い数値であると言える。

3.5 コーパス変更による翻訳精度テスト

コーパスを変更した際に生じる精度の変化を、BLEU によって確認した。「size=512 num_layers=2」のパラメータで、ReLU を用いて、日英法令対訳コーパスと ASPEC を用いて 200,000step 学習を行った。これらのモデルで、そ

れぞれのコーパスの、3.2 の実験と同条件の日本語を翻訳した。この結果を、3.2 の実験と同じく BLEU を計算し集計した。Fig.4 に、コーパスごとの BLEU のグラフを示す。

この結果、ASPEC の方が、日英法令対訳コーパスより BLEU が低くなった。

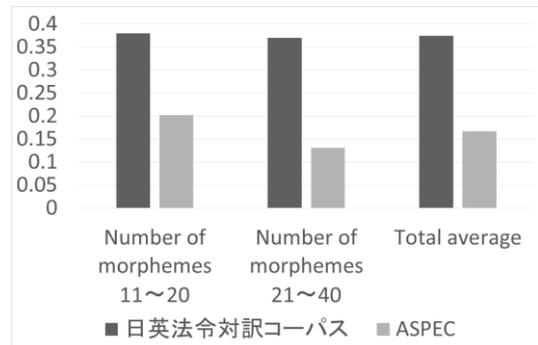


Fig.4 BLEU of each corpus

この結果から、大規模なコーパスを用いる事で、精度は低下すると結論出来る。精度の低下は、二つの原因が考えられる。

一つ目の原因は「コーパスに含まれる表現が多い為、参照訳と異なる似た意味の表現に翻訳する」という点である。

二つ目の原因は、「学習中の perplexity が下がりきっていなかった」という点である。日英法令対訳コーパスを用いた際、全てのモデルの 200,000step 時点で、perplexity は 1.05 以下に下がった。しかし、ASPEC を用いた際は同時時点で 3.00 以上であった為、十分に学習が進んでいない可能性がある。

以上の二つの原因から、翻訳精度が低下したと考える。

4.結論

本研究では、TensorFlow の機械翻訳スクリプトを用いた日本語から英語への機械翻訳を試み、小規模なコーパスで一定の翻訳精度を持つ機械翻訳に成功した。また、機械翻訳が WEB 上で動作するシステムの開発に成功した。更に、GPU を搭載した PC での学習の有用性を示した。

今後は、大規模なコーパスでの精度の確保が課題となる。課題を解決する為に、学習が十分であったかを検証するとともに、更なるモデルの改良が必要である。

参考文献

- (1) ASPEC <http://orchid.kuee.kyoto-u.ac.jp/ASPEC/>
- (2) TensorFlow <https://www.tensorflow.org/>
- (3) 日英法令対訳コーパス <http://www.phontron.com/jaen-law/index-ja.html>
- (4) A.Krizhevsky, et al.: Advances in neural information processing systems, pp1097-1105, 2012.
- (5) I.Sutskever, et al.: Advances in Neural Information Processing System 27, pp.3104-3112, 2014.
- (6) K.Cho, et al.: arXiv:1406.1078, 2014.
- (7) K.Papineni, et al: Proceeding of the 40th Annual Meeting on Association for Computational Linguistics, pp.311-318, 2002.
- (8) S.Hochreiter, et al.: Neural Computation, Vol.9, No.8, pp.1735-1780, 1997.
- (9) Y.LeCun, et al.: Neural networks: Tricks of the trade, pp.9-50, 1998.