

Positional Encoding の文章分類タスクに対する LSTM の代替検討と性能評価

Alternative study and performance evaluation of LSTM for sentence classification task of Positional Encoding

佐藤 充伎
Mitsuki Sato

京都大学理学部
Kyoto University Faculty of Science

Transformer, a neural network model widely used in tasks such as machine translation of natural language processing and sentence classification, obtains position information by Positional Encoding. Using the fact that LSTM is capable of long-term memory, we verified whether LSTM can be used instead of Positional Encoding.

1. はじめに

自然言語処理の機械翻訳や文章分類などのタスクで広く使われるモデルは LSTM や Transformer が有名である。Transformer モデルは位置情報を Positional Encoding により得て、さらに位置情報の依存関係を Attention 機構により学習することにより高性能な文章分類や文章生成タスクを可能にしている。

今回の最終課題を通して私は第 6 回「回帰結合型ニューラルネットワーク」から RNN と LSTM、Transformer の紹介がなされていたことを受け、各アーキテクチャの特性をより深く理解することを目的に文章分類タスクを異なるアーキテクチャ間で比較することにより、それぞれの技術がどれほど優れているか検証した。また LSTM は長期記憶が可能なることを利用して Positional Encoding の代わりに LSTM を使用することができるのか文章分類タスクにおいて検証した。(このアーキテクチャを LSTMTransformer と呼ぶことにする。)

2. 実験方法

映画のレビューの IMDb データセットを使用してレビューが positive な内容か、negative な内容か判定する文章分類タスクに取り組んだ。文章の各単語をトークン化して ID を割り振ったものを訓練データとして使用する。

2.1 RNN

RNN は、前の時間ステップの隠れ状態を次の時間ステップへと「伝播」することで、過去の情報を考慮しながら新たな出力を生成する時系列データやシーケンスデータを扱うことを得意とするモデルである。



2.2 LSTM

LSTM は「ゲート」、「メモリーセル」と呼ばれる構造を持ち、これにより情報を制御します。「忘却ゲート」が古い情報を忘れるかどうかを決定し、「入力ゲート」が新しい情報をどれだけ取り入れるかを決定することで、CEC が長期的な依存関係を保持する。

LSTM は 1 層いれて実験に使用した。

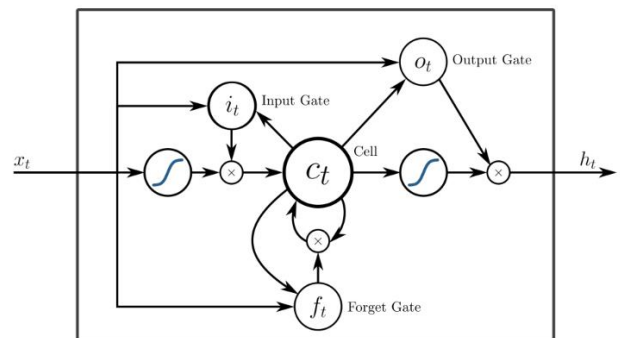


図 3.1 LSTM ユニットの制御フロー

2.3 LSTM(Peephole)

3.2 の各 LSTM のユニットの中に覗き穴ゲートを実装して比較実験に使用した。

2.4 LSTM+Transformer

Transformer のエンコーダ部分は Embedding 層、Positional Encoding 層、Multi-Head Attention ブロックから構成されている。このうち Positional Encoding は文章のトークンの位置情報の獲得のために使用されている。ここに LSTM のユニットで代替して Transformer モデルを使用した。

なお計算コストの都合で LSTM、Transformer ブロックは 1 層ずつ実装した。

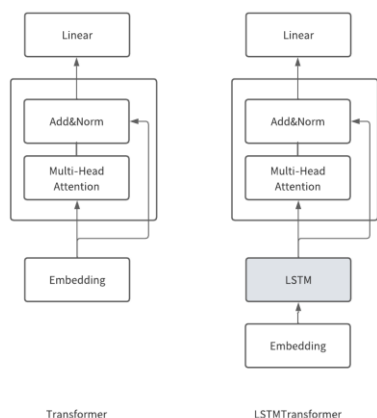


図 3.2 LSTMTransformer モデルの構造

3. 実験結果

epoch数	1	2	3	4	5
RNN	0.5136	0.5696	0.5356	0.6122	0.6398
LSTM	0.468	0.6274	0.6208	0.6692	0.8276
LSTM(Peepphole)	0.4492	0.5452	0.7382	0.824	0.8494
Transformer	0.4348	0.5666	0.7174	0.6932	0.706
LSTMTransformer	0.501	0.725	0.674	0.675	0.65

表 1 各アーキテクチャの F1 スコア比較結果

実験結果は上の表 1 のようになった。これらの数値は各アーキテクチャの小分類タスクにおける F1 スコアの 5 回分の平均をとったものである。これからそれぞれのアーキテクチャ間に焦点を当ててみる。

3.1 LSTM と RNN

LSTM と RNN の F1 スコアを比較した結果、LSTM の方が RNN よりもスコアが高いことが分かった。

RNN と LSTM の比較

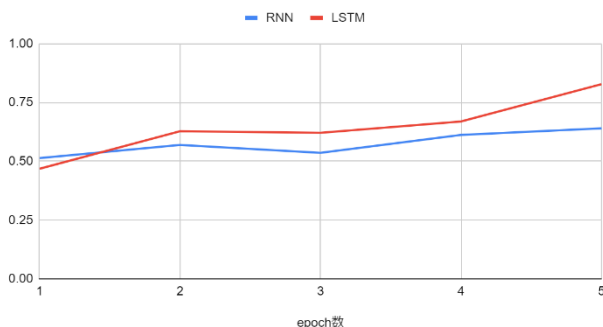


図 4.1 LSTM と RNN のスコア比較

3.2 LSTM と LSTM(Peepphole)

LSTM と LSTM(Peepphole)の F1 スコアを比較した結果、LSTM(Peepphole)の方が LSTM よりもわずかにスコアが高いことが分かった。

LSTM と LSTM(Peepphole)の比較

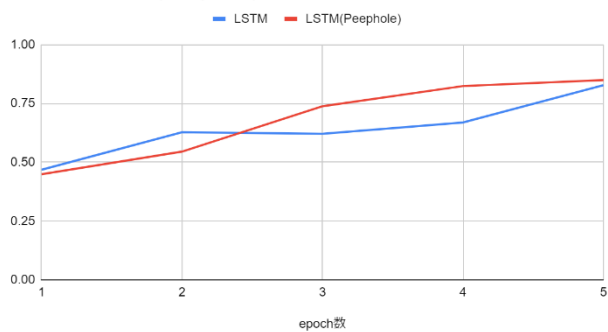


図 4.2 LSTM と LSTM(Peepphole)のスコア比較

3.3 LSTM(Peepphole)と LSTMTransformer

LSTM(Peepphole)と LSTMTransformer の F1 スコアを比較した結果、LSTM(Peepphole)の方が LSTMTransformer よりもスコアが高いことが分かった。

LSTM(Peepphole) と LSTMTransformer

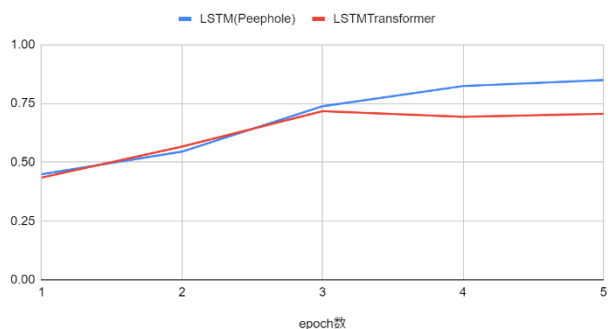


図 4.3 LSTM(Peepphole)と LSTMTransformer のスコア比較

3.4 LSTMTransformer と Transformer

LSTMTransformer と Transformer の F1 スコアを比較した結果、Transformer の方が LSTMTransformer よりもスコアが高いことが分かった。

Transformer と LSTMTransformer

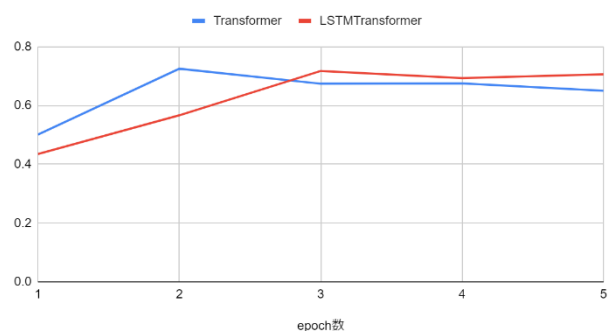


図 4.4 Transformer と LSTMTransformer のスコア比較

3.5 最高 F1 スコアでの比較

最高 F1 スコアで比較した結果以下のグラフのようになった。

アーキテクチャごとの最高スコア

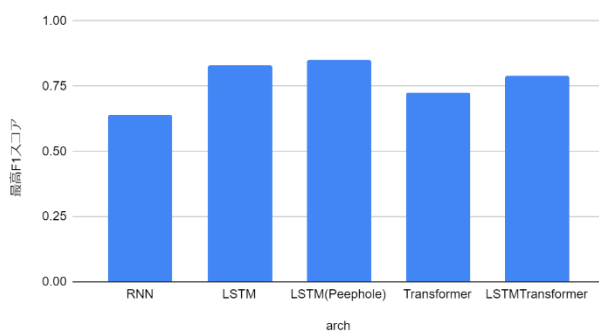


図 4.5 最高 F1 スコアでの各アーキテクチャ間の比較

最高 F1 スコアでも

LSTM(Peepphole)>LSTM>LSTMTransformer の順で性能が高いことがわかった。

4. 考察と結論

今回の実験から文章分類タスクにおいて LSTMTransformer は LSTM と Transformer のいずれにもスコアが及ばず、Positional Encoding は LSTM で代替することはできないことが分かった。

原因として以下の 2 点が考えられる。

- Transformer と LSTM の関係
Transformer では位置情報を Positional Encoding で取得し、Attention 機構で位置情報の依存関係を学習するのに対して Positional Encoding を LSTM で代替すると異なる方法で複数回位置情報の依存関係を学習することになりうまくお互いの特性がかみ合わなかったと考えている。
- 実験設定が合わない。
最適化手法では optimizer まわりの設定や学習率スケジューリングを選ぶことによりさらにスコアの改善が期待できる。