

仮タイトル

1 はじめに

実データとボットの識別。(識別機で生成の評価ができる)

2 要素技術

2.1 Transformer

Transformer[1] は再帰的ニューラルネットワーク (Recurrent Neural Network: RNN)[2] を使わずに, Self-Attention を使って並列計算を可能にするモデルである.

RNN は時系列に沿って順番に計算する構造であるため並列計算ができず, GPU などを使っても計算時間が長いという欠点がある. GPU の性能を十分に活用し, 計算速度を向上させるため, RNN を使用しないことが必要となる.

Self-Attention は, シーケンス内の単語間の関係性に注目する.

Positional Encoding

2.2 GPT-2

Generative Pretrained Transformer 2 (GPT-2)[3] は特定のタスクに特化した数万の教師ありデータでファインチューニングが必要だった GPT, BERT をはじめとしたモデルと異なり, 大規模な言語コーパスで様々なタスクを学習する汎用的なモデルとして設計された.

任意の長さの文章をシンボルにしたシーケンスを (s_1, s_2, \dots, s_n) とする. 同時確率は条件付確率の積に分解される.

$$p(x) = \prod_{i=1}^n p(s_i | s_1, \dots, s_{n-1}) \quad (1)$$

この条件付確率は, Transformer[1] によって効率的に計算できる.

GPT-2 では, $p(\text{output}|\text{input})$ を一般化して, $p(\text{output}|\text{input}, \text{task})$ と表しタスクの種類も学習することで, 特定の問題のためのファインチューニングなしで, 1 つのモデルでも問題を解けるようにしている.

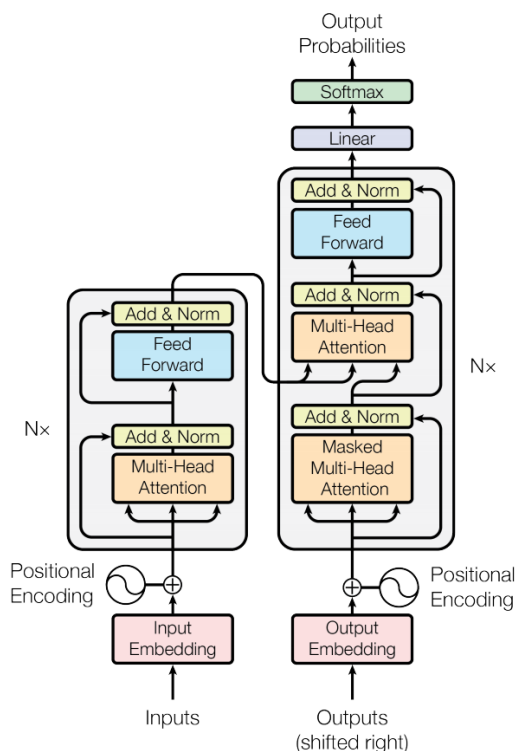


図 1: Transformer

2.3 Byte Pair Encoding

Byte Pair Encoding (BPE) は高頻度の単語は単語全体を辞書に登録し, 低頻度の単語は文字単位に分割する言語モデルにおける未知語処理手法である.

言語モデルは語彙サイズをハイパーパラメータとして設定するため, ニューラルモデルに扱われない未知語が存在する. 未知語処理として, $\langle unk \rangle$ などの特殊トークンに置き換える方法と, 単語をより細かく分割しサブワードにすることで語彙を少なくする方法がある.

BPE はデータ圧縮手法を言語モデルに応用した. 文字単位に単語を分解し, 2 文字のペアの中で高頻度の要素を結合して 1 つのサブワードとする手順を繰り返すことで, 単語を接頭辞や接尾辞などの意味のある単位に分解できる.

表 1: GPT-2 の解けるタスク

AutomaticSpeechRecognition
Conversational
FeatureExtraction
FillMask
ImageClassification
QuestionAnswering
Summarization
TextClassification
TextGeneration
TokenClassification
Translation
ZeroShotClassification

2.4 BERT

BERT [4] は, 2018 年に Google が発表した汎用言語モデルである. 複数の双方向 Transformer を用いることで文脈を考慮することができるモデルとされている. 各タスクに対してファインチューニングをすることでさまざまなタスクに柔軟に対応することができる. 事前学習には入力の一部の単語を “[MASK]” に置き換えてその元単語を予測するように訓練するタスク (Masked Language Model) と 2 文を入力としてその連続性を識別するように訓練するタスク (Next Sentence Prediction) が用意されている. 本研究の実験では, 京都大学によって公開されている BERT 日本語 Pretrained モデルを使用した.

3 実験

3.1 対話タスク

図 2 に対話例を示した. 英語の場合, 対話を自然に続けられることを確認した.

参考文献

- [1] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *CoRR*, abs/1706.03762, 2017.

```
>> User:Hello
DialogPT: Hiya
>> User:What your name?
DialogPT: I'm a girl
>> User:How is the weather today?
DialogPT: It's nice
>> User:It sounds good
DialogPT: I'm glad
>> User:Good bye
DialogPT: I'll see you around
```

図 2: 対話例

- [2] Tomas Mikolov, Martin Karafiat, Lukas Burget, Jan Cernocky, and Sanjeev Khudanpur. Recurrent neural network based language model. In *Eleventh annual conference of the international speech communication association*, 2010.
- [3] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9, 2019.
- [4] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.