近世・近代・現代日本語テキストに対する場所参照表現抽出

片山歩希 [1]　東山翔平 [2,1,3]　大内啓樹 [1,6,3]　坂井優介 [1]
竹内綾乃 [4]　坂東諒 [3]　橋本雄太 [5]　小木曽智信 [3]　渡辺太郎 [1]
[1] 奈良先端科学技術大学院大学　[2] 情報通信研究機構　[3] 国立国語研究所
[4] 国文学研究資料館　[5] 国立歴史民俗博物館　[6] 理化学研究所
{katayama.ayuki.kc1,hiroki.ouchi,sakai.yusuke.sr9,taro}@is.naist.jp,
shohei.higashiyama@nict.go.jp,　takeuchi.ayano@nijl.ac.jp,
yhashimoto@rekihaku.ac.jp,　{r-bando,togiso}@ninjal.ac.jp

概要

歴史的テキストからの場所参照表現の抽出は，大規模な史料に対する人文学的分析を支援するための基礎技術として重要である．本研究では，近代・近世の日本語テキストを用いたデータセットを構築し，これら歴史的テキストに対する Transformer 言語モデルの抽出精度を調査した．実験から，現代語ラベル付きデータの活用の有効性を確認した一方で，歴史的テキストへの適応には，さらなるモデルの改善が必要であることも示された．

1　はじめに

各時代の出来事や人間の活動が記録された歴史的テキストは，人類史や自然史をより深く理解するための史料として重要である．地理的な観点から見ると，歴史的テキストには，歴史的地名・施設名などの場所参照表現（Location Referring Expressions; LRE）が含まれ，その場所に関連する人，物，出来事などとともに記述されていることが多い．たとえば，紀行文では筆者が訪れた場所での体験が記述され，災害記録では被災地域，被害の規模，被災者の状況などが記述される．計算機によって，こうした地理的情報の高精度な自動抽出・構造化が実現されると，「遠読」とも呼ばれる大規模な歴史的テキストの横断的分析など，人文学研究における分析作業を支援することが可能になる．

本研究では，計算機による地理的テキスト解析の基本的なステップとして，歴史的日本語テキストからの LRE 抽出に取り組む．例として，“名取川渡りて仙台に入る” という入力文に対しては，“名取川” および “仙台” を抽出することが目標となる．

評価対象のモデルとして，本タスクにおいて高い性能が期待できる Transformer [1] 言語モデルを用いた．具体的には，Masked Language Model（MLM）である日本語 BERT [2] モデルと，Causal Language Model（CLM）である Llama-3-ELYZA-JP-8B（ELYZA）[3] について，fine-tuning を実施し，評価を行った．

モデルの性能評価実験には，次の 4 つのデータセットを使用した．我々が LRE アノテーションを施した近世紀行文「おくのほそ道」および近代紀行文「突貫紀行」[1)]，近世の災害記録からなる「みんなで翻刻データ」[4]，現代紀行文（旅行記）の「地球の歩き方旅行記」[5, 6, 7] である．

近世・近代に加え，現代のテキストを使用する理由は 2 つある．(i) 異なる時代のテキストに対する抽出精度の違いと，(ii) 歴史的テキストに対する抽出精度向上に現代語テキストを活用することの有効性を調査するためである．日本語 LRE 抽出において，近世・近代・現代の 3 時代を横断した評価・分析を行った研究は，我々が知る限り本研究が初である．

実験結果から，以下の知見が得られた．

- BERT，ELYZA 両モデルは，現代のテキストでは高い精度を達成したものの（最大 F1 値 0.886），歴史的テキストでは低〜中程度の精度にとどまった（最大 F1 値 0.434–0.682）．
- 現代および近代テキストでは ELYZA がより高い精度を示し，近世テキストでは BERT，ELYZA で概ね同等の精度となった．
- 近世のみでなく現代のラベル付きデータも fine-tuning に用いることで，BERT，ELYZA とも各歴史的テキストでの精度が向上した．

1）アノテーションデータセットは次の URL で公開する：https://github.com/naist-nlp/historical-travelogues.

2 関連研究

場所参照表現（LRE）抽出は、固有表現抽出（Named Entity Recognition; NER）[8] の特殊なケースに相当する．LRE が指す場所に対応する地理座標を特定するタスクはジオコーディングと呼ばれ，LRE 抽出は，ジオコーディングと合わせてジオパージング [9] として取り組まれることも多い．

言語資源　歴史的テキストを用いた LRE アノテーションデータとして，英語ニュース記事 [10]，英語旅行記 [11, 12]，フランス語文学テキスト [13]，中国語歴史書 [14] などのデータセットが構築され，機械学習システムの性能が評価されてきた．日本語については，地震史料集テキストデータベース [15] や，みんなで翻刻 [4] など，歴史的な災害記録テキストに対して，LRE とその地理座標を人手アノテーションする取り組みが行われている．

システム性能評価　LRE を含む固有表現の認識のための様々な手法について，歴史的テキストに対する性能が調査されてきた [16]．最近の研究では，事前学習 Transformer 言語モデルを用いたものがある．Labsch ら [17] は，ドイツ語歴史的新聞記事に対する NER において，BERT のための学習戦略を調査した．同研究では，古典語ラベル付きデータでの fine-tuning の前に，大規模な古典語ラベルなしデータと，現代語ラベル付きデータの両方で事前学習したモデルが最も高精度であったことが示されている．Tang ら [14] は，古代中国の歴史的文書に対する NER において，歴史的テキストで事前学習した MLM と，オープンおよびクローズドな現代語事前学習 CLM を評価し，MLM がより高精度であることを示した．

3 実験設定とデータセット

LRE 抽出タスクにおいて，大規模な現代語テキストで事前学習された言語モデルを，歴史的テキスト（古典語テキスト）に適応させるための実験を行う．そこで，後述する 4 データセットを用いた学習・評価シナリオを設定した．学習シナリオとして，

1. 現代語のラベル付きデータでの fine-tuning
2. 古典語のラベル付きデータでの fine-tuning
3. 現代語・古典語両方のラベル付きデータでの fine-tuning

の 3 種類を設定し，各学習済みモデルについて，現代語・古典語の各評価データで評価する．これらの評価を通じて，MLM と CLM の精度を比較する．

表 1　各データセットの記述統計（文数，LRE 数）

Data	Train		Dev		Test	
	Sent.	LRE	Sent.	LRE	Sent.	LRE
Arukikata	6,516	3,102	601	260	5,156	2,166
Minna	1,800	9,585	101	178	476	2,408
Hosomichi	–	–	–	–	523	242
Tokkan	–	–	–	–	180	117

3.1 データセット

既存の現代語データセット 1 件（Arukikata）と古典語データセット 1 件（Minna）に加え，新たに古典語データセット 2 件（Hosomichi，Tokkan）を作成した．各データセットの記述統計を表 1 に示す．

3.1.1 地球の歩き方旅行記（Arukikata）

現代語テキストとして，現代語の日本国内旅行記に人手で LRE がアノテーションされたデータセットである ATD-MCL [7] を使用した．`LOC-NAME`（地名）または `FAC-NAME`（施設名）のラベルが付与された固有名の LRE を対象とし，ラベルを `LOCATION` に統一した上で，実験に用いた．訓練・開発・テストセットの分割は，文献 [7] に従った．

3.1.2 みんなで翻刻（Minna）

1 つ目の歴史的テキストとして，みんなで翻刻 [4] プロジェクト[2]で構築されたアノテーションデータを使用した．同データは，1800 年代前後の近世日本の災害記録のテキストからなり，「日時」，「場所」，「現象・被害」，「人物」を表す表現が人手アノテーションされている[3]．データの前処理は次のように行った．まず，原文テキスト[4]に，「場所」のアノテーション情報[5]を `LOCATION` ラベルの LRE として統合した後，LRE が 1 件以上出現する txt ファイル（原書の見開き 1 ページに相当）全体を選択した．次に，それらの各 txt ファイルを 50 文字ごとにセグメントに分割し，各セグメントを文とみなして学習・評価の入力単位とした[6]．次に，文集合全体からランダムに 80%の文を抽出し，そのうち約 95%，

2) https://honkoku.org/index.html
3) https://wiki.honkoku.org/doku.php?id=annotation-top
4) https://github.com/yuta1984/honkoku-data/tree/master/v1
5) https://ansei2.vercel.app/api/annotations?type=location
6) セグメントを跨ぐ LRE は，LRE でないものと扱った．

5%をそれぞれ訓練，開発セットとし，残りの 20%をテストセットとした（つまり，3 セットの文数の比率は概ね 76:4:20 となる）．

本データのテキストの特徴として，地震の場所や被害の描写を列挙するスタイルで書かれたものが多い点が挙げられる．例：“小石川御門内よりするが臺 小川丁 筋違御門迄少〻破損”（下線部は LRE を示す）．

3.1.3 おくのほそ道（Hosomichi）

2 つ目の歴史的テキストとして，Wikisource で公開されている「おくのほそ道」のテキスト[7]を使用し，LRE アノテーションデータセットを作成・使用した．「おくのほそ道」は，1690 年代頃に松尾芭蕉によって書かれたもので，同時代を代表する歴史的紀行文の一つである．「おくのほそ道」は人間の地理的移動に焦点を当てた文学テキストであり，実用上の目的から地理的自然現象が記録された文書である「みんなで翻刻」のテキストと比較すると，ともに近い時代に書かれたテキストでありながら，異なる特徴を持つと考えられる．

アノテーション作業は，日本古典文学作品のコーパス構築経験のある著者 2 名により，文境界付与，LRE 付与という手順で行った[8]．なお，本データセットはサイズが小さいため，未知の近世テキストデータとして評価にのみ使用した[9]．

3.1.4 突貫紀行（Tokkan）

3 つ目の歴史的テキストとして，青空文庫で公開されている「突貫紀行」のテキスト[10]を使用し，LRE アノテーションデータセットを作成・使用した．「突貫紀行」は，明治時代に幸田露伴によって書かれ，1893 年に発表されたもので，同時代を代表する歴史的紀行文の一つである．「突貫紀行」は，「おくのほそ道」と同様に文学的な紀行文であるが，より現代に近い歴史的テキストである．

「おくのほそ道」と同様に，著者 2 名により文境界・LRE のアノテーションを行った．また，本データも，未知の近代テキストとして評価にのみ使用した．

7） https://ja.wikisource.org/wiki/%E3%81%8A%E3%81%8F%E3%81%AE%E3%81%BB%E3%81%9D%E9%81%93

8） 基準の概要について付録 A に簡潔に記す．

9） 事前学習モデルの学習に生テキストが使用された可能性はあるが，ラベルについては未知である．

10） https://www.aozora.gr.jp/cards/000051/files/830_14079.html

```
指示:\n 次の紀行文の文章を読んで、地名・施設名についての質問に回答してください。\n
文章:\n{TEXT}\n
質問:\n 旅行記から地名・施設名を抽出してください。\n
出力形式:\n 地名・施設名をひとつも抽出しなかった場合は「なし」と出力してください。複数の地名・施設名を抽出した場合は「@@」で区切って出力してください。\n
回答:\n
```

図 1 CLM 用プロンプト．“{TEXT}” には実際のテキストが挿入される．一部の改行文字は “\n” で示している．

3.2 言語モデル

大規模な現代語日本語テキストで事前学習された，MLM および CLM を評価に用いた．各モデルの fine-tuning におけるハイパーパラメタの設定は付録 B に示す．

BERT MLM として，日本語文字単位の BERT [2] 事前学習モデル[11]（Large，パラメタ数 340M）を用いた．各文字トークンにラベル（B-LOCATION, I-LOCATION, O）を割り当てる系列ラベリングとして LRE 抽出を行うため，ラベル分類用の全結合層を追加し，Softmax 交差エントロピー損失を用いて fine-tuning を行った．

ELYZA CLM として，Llama-3 [18] に日本語データでの継続的事前学習を施した Llama-3-ELYZA-JP-8B [3][12]（ELYZA モデルと呼ぶ）を使用した．図 1 の形式のプロンプトを使用し，「回答:\n」に続くテキストを生成するよう，QLoRA [19] を用いて fine-tuning を行った．

4 実験結果

各学習シナリオ（§3 冒頭で述べた 1.〜3. の学習データ）において，BERT，ELYZA モデルの異なる乱数シードでの学習をそれぞれ 3 回実行し，各学習において開発セットでの F1 値が最良であったモデルチェックポイントを保存し，評価した．各テストセットでの抽出精度（F1 値）を表 2 に示す．以降，§4.1–4.3 でデータ別での各モデル・学習シナリオについての結果を議論した後，§4.4 でデータ横断での結果を議論する．

4.1 現代旅行記での評価結果

Arukikata テストセットでは，BERT，ELYZA とも Arukikata 学習セットを含むデータで学習した場合に，高い精度（F1 値 0.852–0.886）を達成した．

11） https://huggingface.co/tohoku-nlp/bert-large-japanese-char-v2

12） https://huggingface.co/elyza/Llama-3-ELYZA-JP-8B

表 2 各モデルの LRE 抽出精度．学習データの “A” は Arukikata データ，“M” は Minna データを指す．下線は同一モデル内，太字はモデル横断での最高精度を示す．

Model	Train	Test			
		Aruki	Minna	Hoso	Tokkan
BERT	A	0.843	0.232	0.350	0.494
	M	0.223	0.657	0.329	0.496
	A+M	0.830	**0.682**	0.417	0.524
ELYZA	A	**0.886**	0.253	0.375	0.649
	M	0.509	0.664	0.352	0.497
	A+M	0.880	0.674	**0.434**	**0.670**

Arukikata 学習セットに加えて Minna 学習セットを用いたケースでは，多少の変化は見られたものの小さい影響であった．主な理由として，Arukikata データに対する Minna データのスタイル（内容および時代）の違いが大きい点が挙げられる．

また，モデル間の比較では，いずれの学習シナリオにおいても ELYZA が BERT を上回り，特に Minna 学習セットのみで学習したケースでの差が顕著である．これは，ELYZA が，高性能なベースモデル Llama-3 の能力を引き継ぎつつ，継続事前学習によって日本や現代日本語に関する高度な知識・読解能力を有したことを示唆している[13]．

4.2 近世地震記録での評価結果

Minna テストセットでは，両モデルとも，Minna 学習セットのみの場合に比べ，Arukikata 学習セットを追加で用いた場合に，やや精度が向上した．Arukikata データには日本全国の様々な地名が含まれており，それが Minna データ（江戸周辺の地名に偏っている）での抽出に対して有効であった可能性がある．ただし，前述したような両データのスタイルの違いからか，Arukikata 学習セットのみでの学習では非常に低い精度にとどまっている．

モデル間の比較としては，どの学習シナリオでも両モデルはほぼ同等の精度を示した．

4.3 近世・近代紀行文での評価結果

Hosomichi テストセット，Tokkan テストセットでは，両モデルとも，Minna，Arukikata の一方よりも両方の学習セットを用いた場合に顕著に精度が向上し，最高精度を示した．この結果は，近世・近代紀行文からの抽出において，現代の旅行者の記録である Arukikata データと，近世の地震の記録である Minna データは，異なる特徴を有しつつも共に有用であり，相補的な効果があったことを示している．

モデル間の比較では，Hosomichi テストセットでは両モデルで同水準の精度であったが，Tokkan テストセットでは ELYZA の方が概ね高精度であった．Tokkan データの方が現代語に近いテキストであることから，この結果も，ELYZA の高い日本語知識・能力を示唆している．

4.4 総括

データ横断での全体的な傾向を述べる．最も高精度となる傾向が見られた 2 つの学習セットを用いた場合に注目すると，両モデルとも，Arukikata テストセットにおいて高い精度（F1 値 0.8 以上）を示し，続いて Minna および Tokkan において中程度の精度（F1 値 0.5–0.7）を示し，Hosomichi では最も低い精度（F1 値 0.5 未満）を示した．この点は，各言語モデルが現代日本語テキストで事前学習されており，Minna ではドメイン内データでの学習を行っている点，Tokkan では比較的現代に近いテキストである点を踏まえると，直感的な結果である．

各歴史的テキストでの抽出精度を高めるためには，個別にドメイン内ラベル付きデータを作成（増量）して学習に用いる他，Labusch [17] らと同様の戦略として，大規模な古典語のラベルなしデータや，指示学習データなどを事前学習に使用し，古典語の知識・処理能力を高めることが必要と考えられる．なお，現代，近代のテキストでは，BERT よりも ELYZA が高い精度を達成したものの，現代から離れた時代のテキストにおいてどのようなモデルが高精度となるかは探求の余地がある．

5 おわりに

本研究では，我々が構築した近代・近世紀行文の場所参照表現（LRE）アノテーションデータセットとともに，既存の現代および近世テキストデータセットを使用し，各時代のテキストに対する日本語言語モデルの LRE 抽出精度を調査した．実験から，現代語・古典語両方のラベル付きデータで fine-tuning することの有効性を確認した．

今後の展望として，(1) 古典語ラベルなしデータの活用も含め，歴史的テキストを高精度に解析可能とするモデル学習方法の探求，(2) 歴史的テキストに対するジオコーディングのためのデータおよび手法の構築・評価などに取り組む予定である．

13） 文献 [20] では，Llama-2 からの継続事前学習モデルである Swallow-7b-hf は BERT に劣る結果を示している．

謝辞

本研究は JSPS 科研費 JP23K24904 と国立国語研究所 共同利用型共同研究 (B)「歴史的日本語資料のためのジオパージング」および基幹研究プロジェクト「開かれた共同構築環境による通時コーパスの拡張」の助成を受けたものです.

参考文献

[1] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, L ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In **Advances in Neural Information Processing Systems**, Vol. 30. Curran Associates, Inc., 2017.

[2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional Transformers for language understanding. In **Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)**, pp. 4171–4186. Association for Computational Linguistics, June 2019.

[3] Masato Hirakawa, Shintaro Horie, Tomoaki Nakamura, Daisuke Oba, Sam Passaglia, and Akira Sasaki. elyza/Llama-3-ELYZA-JP-8B. https://huggingface.co/elyza/Llama-3-ELYZA-JP-8B, 2024.

[4] 橋本雄太. 歴史災害資料のマークアップシステムの試作. 研究報告人文科学とコンピュータ, Vol. 2023-CH-131, No. 2, pp. 1–6, 2023.

[5] 株式会社地球の歩き方. 地球の歩き方旅行記データセット. 国立情報学研究所 情報学研究データリポジトリ. https://doi.org/10.32130/idr.18.1, 2022.

[6] Hiroki Ouchi, Hiroyuki Shindo, Shoko Wakamiya, Yuki Matsuda, Naoya Inoue, Shohei Higashiyama, Satoshi Nakamura, and Taro Watanabe. Arukikata travelogue dataset. arXiv:2305.11444, 2023.

[7] Shohei Higashiyama, Hiroki Ouchi, Hiroki Teranishi, Hiroyuki Otomo, Yusuke Ide, Aitaro Yamamoto, Hiroyuki Shindo, Yuki Matsuda, Shoko Wakamiya, Naoya Inoue, Ikuya Yamada, and Taro Watanabe. Arukikata travelogue dataset with geographic entity mention, coreference, and link annotation. In **Findings of the ACL: EACL 2024**, pp. 513–532, March 2024.

[8] David Nadeau and Satoshi Sekine. A survey of named entity recognition and classification. **Lingvisticae Investigationes**, Vol. 30, No. 1, pp. 3–26, 2007.

[9] Milan Gritta, Mohammad Taher Pilehvar, and Nigel Collier. A pragmatic guide to geoparsing evaluation: Toponyms, named entity recognition and pragmatics. **Language Resources and Evaluation**, Vol. 54, pp. 683–712, 2020.

[10] Mariona Coll Ardanuy, David Beavan, Kaspar Beelen, Kasra Hosseini, Jon Lawrence, Katherine McDonough, Federico Nanni, Daniel van Strien, and Daniel C. S. Wilson. A dataset for toponym resolution in nineteenth-century English newspapers. **Journal of Open Humanities Data**, Jan 2022.

[11] Paul Rayson, Alex Reinhold, James Butler, Chris Donaldson, Ian Gregory, and Joanna Taylor. A deeply annotated testbed for geographical text analysis: The corpus of lake district writing. In **Proceedings of the 1st ACM SIGSPATIAL Workshop on Geospatial Humanities**, p. 9–15. Association for Computing Machinery, 2017.

[12] Rachele Sprugnoli, et al. Arretium or arezzo? A neural approach to the identification of place names in historical texts. In **Proceedings of the Fifth Italian Conference on Computational Linguistics**. CEUR-WS, 2018.

[13] Eleni Kogkitsidou and Philippe Gambette. Normalisation of 16th and 17th century texts in French and geographical named entity recognition. In **Proceedings of the 4th ACM SIGSPATIAL Workshop on Geospatial Humanities**, p. 28–34. Association for Computing Machinery, 2020.

[14] Xuemei Tang, Qi Su, Jun Wang, and Zekun Deng. CHisIEC: An information extraction corpus for Ancient Chinese history. In **Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)**, pp. 3192–3202. ELRA and ICCL, May 2024.

[15] Yasuyuki Kano and Junzo Ohmua. Integration of geographical information into the data of materials for the history of Japanese earthquakes. **IPSJ SIG Computers and the Humanities Technical Report**, Vol. 2023-CH-131, No. 3, pp. 1–3, 2023.

[16] Maud Ehrmann, Ahmed Hamdi, Elvys Linhares Pontes, Matteo Romanello, and Antoine Doucet. Named entity recognition and classification in historical documents: A survey. **ACM Comput. Surv.**, Vol. 56, No. 2, sep 2023.

[17] Kai Labusch, Preußischer Kulturbesitz, Clemens Neudecker, and David Zellhöfer. BERT for named entity recognition in contemporary and historical German. In **Proceedings of the 15th Conference on Natural Language Processing (KONVENS 2019)**, pp. 9–11, 2019.

[18] Aaron Grattafiori et al. The Llama 3 herd of models. arXiv:2407.21783, 2024.

[19] Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. QLoRA: Efficient finetuning of quantized LLMs. In **Thirty-seventh Conference on Neural Information Processing Systems**, 2023.

[20] Ayuki Katayama, Yusuke Sakai, Shohei Higashiyama, Hiroki Ouchi, Ayano Takeuchi, Ryo Bando, Yuta Hashimoto, Toshinobu Ogiso, and Taro Watanabe. Evaluating language models in location referring expression extraction from early modern and contemporary Japanese texts. In **Proceedings of the 4th International Conference on Natural Language Processing for Digital Humanities**, pp. 331–338. Association for Computational Linguistics, November 2024.

A　アノテーション基準

「おくのほそ道」および「突貫紀行」へのLREアノテーション作業では，表3に示すラベルを定義し，該当する表現にアノテーションを行った．個々の表現をLREとみなすかどうかの基準については，文献[7]のアノテーション基準を参考にしつつ，各歴史的テキストに適した基準に更新した．なお，実験では，LOC-NAMEおよびFAC-NAMEラベルをLOCATIONラベルに統一し，同ラベルの事例のみ対象とした．

表3　LREアノテーションにおけるラベルセット

Label	Description
LOC-{NAME,GENERAL}	地名の固有名，一般名詞句
FAC-{NAME,GENERAL}	施設名の固有名，一般名詞句
VEHICLE	乗り物名（固有名か否か区別なし）
DEICTIC	上記いずれかを指す指示表現等

B　モデルのハイパーパラメタ

表4および表5に，BERT-Largeおよび Llama-3-ELYZA-JP-8Bモデルの fine-tuningにおいて使用したハイパーパラメタの設定をそれぞれ示す．

表4　BERT-Largeに関するハイパーパラメタ

Hyper-parameter	Value
training epochs	20
batch size	32
learning rate	1e-5
lr scheduler type	linear
warmup ratio	0.1
gradient norm clipping threshold	1.0
optimizer	AdamW

表5　Llama-3-ELYZA-JP-8Bに関するハイパーパラメタ

Hyper-parameter	Value
training epochs	10
batch size	8
learning rate	5e-5
lr scheduler type	linear
optimizer	paged_adamw_8bit
quant_method	BITS_AND_BYTES
load_in_4bit	True
bnb_4bit_use_double_quant	True
bnb_4bit_quant_type	nf4
bnb_4bit_compute_dtype	float16
lora_alpha	16
lora_dropout	0.1
bottleneck_r	64
torch_dtype	float16