

ゼミ資料

1 今週やったこと

- RoBERTa モデルで比較
- PLaMo だけでラベルを推定できるか

2 RoBERTa 事前学習済モデルで実験

後期発表では東北 BERT のみを扱っていたため、別モデルでの精度が見たいという目的のもと RoBERTa 事前学習済モデル ("rinna/japanese-roberta-base") に切り替えて実行しようと思ったが、エラーを解決できなかった。東北 BERT との仕様違いの所で実装ミスがあるはず... 早稲田¹の方で今実験回しています。

3 PLaMo だけでラベルを推定できるか

PLaMo API ベータ版を用いて、

- 記事のタイトルと本文情報
- 各記事ラベルの情報と簡潔な説明文
- 出力として想定される JSON Schema

を与えた時に、単純にラベルを識別できるかについてテストデータセットのみに対して調べた。(LLM のファインチューニングではなく、単に質問に答えられるか。) この時、温度パラメータの値は 0.0 とした。図 1 に JSON Schema, 図 2 にプロンプトの概要を示す。

まず、殆どのデータに対して

```
{"results": {  
  "1st": 3,  
  "2nd": 0,  
  "3rd": 7  
}
```

のような正しい JSON 形式での回答を得ることができたが、全 768 データの中で 3 件のデータでは 0 から 8 以外の数字 (例: 269, 14.4) を含んでいた。複数のデータでは正しい JSON に加えて補足文 (「以上が結果です。」のような文章) も出力されていた。

表 1 にテストデータに対する識別結果を示す。比較として 1 位が真のラベルの場合のみ正解とした場合と、上位 3 位までに真のラベルがあれば正解とした場合を示す。

表 1: 実験結果

	Acc	F1
後期発表	0.9687	0.9683
1 位のみ参照	0.5135	0.4488
3 位まで参照	0.7989	0.7976

```
schema = {  
  "type": "object",  
  "results": {  
    "1st": {"type": "int", "description": "最もふさわしいラベル番号"},  
    "2nd": {"type": "int", "description": "2番目にふさわしいラベル番号"},  
    "3rd": {"type": "int", "description": "3番目にふさわしいラベル番号"}  
  }  
}
```

図 1: JSON Schema

また、図 3 4 に、混合行列を示す。上位 1 つだけ見た時点で、特定のクラスに関しては高い精度を示しているが、独女 と Peachy など似たラベルに関しては片方に集中して分類されていることが分かった。

4 次やること

- LLM を用いてデータの重要語を上位 n 語抽出して、そのデータのみで学習をし、精度がどれだけ落ちるか。
- LLM を用いてデータの重要文を上位 m 文抽出して、そのデータのみで学習をし、精度がどれだけ落ちるか。
- LLM を用いてデータの要約文の出力形式を短文 m 文で厳格にし (学習データの簡素化)、そのデータを活用して学習をし、精度がどれだけ落ちるか。

以上。

¹<https://huggingface.co/nlp-waseda/roberta-base-japanese>

5 追記

早稲田 RoBERTa 事前学習済モデル の結果について, Acc 0.8994, F1 値 0.8986 となった. モデル名の所を変えただけのため, 学習曲線を見てみるとまだ学習途中のように思えた. また, p , q , r の学習推移において, 後期発表の場合よりも r を重視しており, p の値が途中から下がり始めている新しい傾向が見られた. 学習パラメータなどの調製が必要.

```

messages=[
  {"role": "system", "content": "あなたは優秀なAIアシスタントです。"},
  {"role": "user", "content": '「{0}」というタイトルの記事の本文が与えられるので、'.format(batch['title']][i]}),
  {"role": "user", "content": "最もこの記事にふさわしい記事ラベルを次の9つから上位3つを選び、選ばれたラベル番号を以下のJSON Schemaに厳密に従ってJSONで返答してください。¥n"},
  {"role": "user", "content": "必ず result という key を持ち、その中には子要素として 1st, 2nd, 3rd を key として持つJSONを出力してください。"},
  {"role": "user", "content": "最終応答は、¥" {¥"で始まり¥" ¥"で終わる、または¥" {¥"で始まり¥" ¥"で終わるJSONのみを出力し、JSON以外の文字は一切応答に含めないでください。¥n"},
  {"role": "user", "content": "JSON Schema:¥n{0}".format(schema)},
  {"role": "user", "content": "記事ラベルとラベル番号は以下のように対応しています。¥n"},
  {"role": "user", "content": "ラベル番号 0: 独女通信（主に独身女性をターゲットにしたライフスタイルや恋愛、ファッション、自己啓発に関する記事）¥n"},
  {"role": "user", "content": "ラベル番号 1: IT ライフハック（主にITやテクノロジー関連の話題やライフハックに関する記事）¥n"},
  {"role": "user", "content": "ラベル番号 2: 家電チャンネル（家電製品やガジェットに関する情報、レビュー記事）¥n"},
  {"role": "user", "content": "ラベル番号 3: livedoor HOMME（主に男性をターゲットとしたライフスタイル、ファッション、趣味に関する記事）¥n"},
  {"role": "user", "content": "ラベル番号 4: MOVIE ENTER（映画やエンターテインメントに関するニュースやレビュー）¥n"},
  {"role": "user", "content": "ラベル番号 5: Peachy（女性向けのライフスタイル全般、ビューティー、恋愛に関する記事）¥n"},
  {"role": "user", "content": "ラベル番号 6: エスマックス（スマートフォンを中心としたモバイル関連情報に関する記事）¥n"},
  {"role": "user", "content": "ラベル番号 7: Sports Watch（スポーツ全般のニュースや情報、特に試合の結果や選手の動向について詳しく報じる記事）¥n"},
  {"role": "user", "content": "ラベル番号 8: トピックニュース（国内外の社会的なニュースやトピックに関する記事）¥n"},

  {"role": "user", "content": "記事の本文:¥n{0}".format(body)},
  {"role": "user", "content": "JSON:"}]

```

図 2: プロンプト

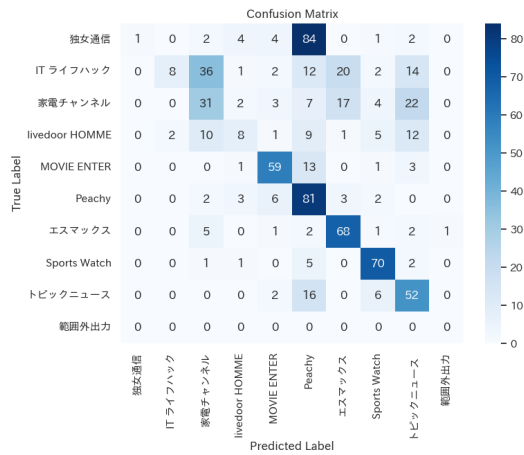


図 3: 混合行列 (上位 1 位まで参照)

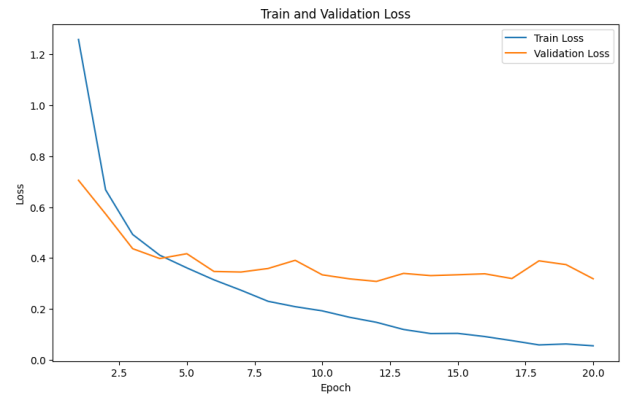


図 6: RoBERTa loss

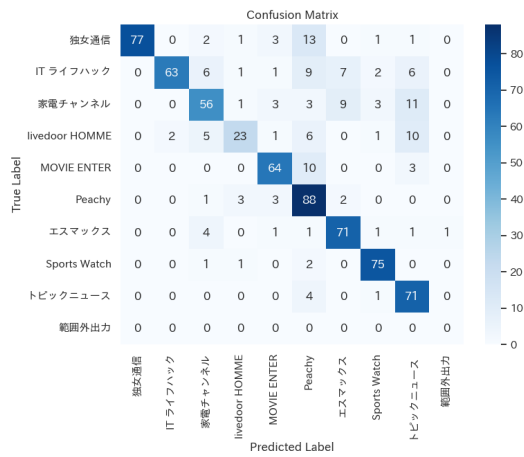


図 4: 混合行列 (上位 1 位まで参照)

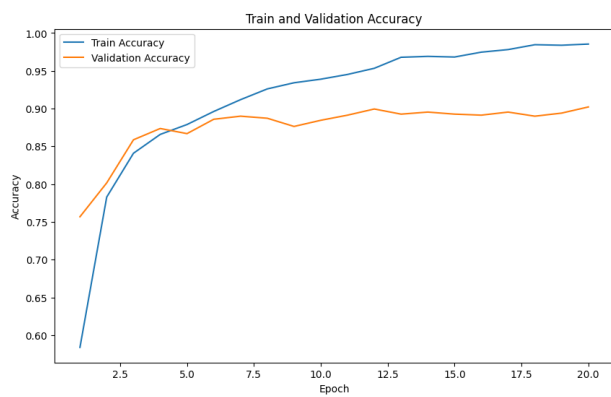


図 5: RoBERTa Acc

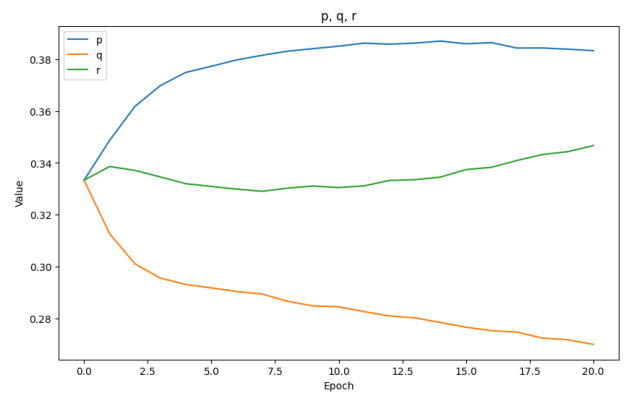


図 7: RoBERTa p,q,r