

半教師あり学習を用いた動機付け面接における MITI コードの予測精度改善に向けた検討

上間 翔 當間 愛晃

琉球大学工学部工学科知能情報コース

MITI とは動機付け面接を評価する行動的コード化システムのことである。しかし、この MITI をコーディングするためには コストがかかる。よって本研究では、Bert を用いて分類を自動で行うことでコストを低減することを目的とする。準備実験では 80% を超える精度が出たが、データ不足が読み取れるような結果であった。今後は、データ数不足を改善することで全ての MITI コードに対応できるよう改善することを試みる。

1 目的と背景

人は置かれている状況に対して何らかの不満や問題があった際に、その状況から抜け出そうと何らかの行動を起こすことが考えられる。しかし現状を変えたいと思うも行動を起こすことができない。動機付け面接とは、そのような人を対象に行われるカウンセリングのアプローチである。面接者が患者に直接的に指示を行うのではなく、患者の気持ちや状況を聞き出し、変わりたいと言う発言を強くすることによって、患者の変化を促すというものである[1]。

MITI (Motivational Interviewing Treatment Integrity) とは動機付け面接において臨床家が面接をうまく使いこなしているかを評価する、行動的コード化システムのことである。評価する対象としては関わる、引き出す、計画するといった 動機付け面接内の構成要素のプロセスを評価する[2]。しかし、この MITI をコーディングするには訓練などでコストがかかってしまう。そこで、この MITI のコーディングを機械学習によって自動化することによって、コストを低減した上でコーディングを行い、面接者への評価を素早く行うことを目的とする。具体的なコードを表 1 に示す。

2 関連研究

動機付け面接における MITI のコード推定は津波が行なっている。津波が行なった研究[3]では、労働安全衛生総合研究所・過労死等防止調査研究センターの木内らによってコード化された日本語のデータセットを用いていた。具体的には、単純な聞き返し、複雑な聞き返し、開かれた質問、閉じた質問、是認、要約、チェンジトーク、維持トーク、その他の計 9 つのラベルでコーディングしている。

表 1 本実験で使った MITI コード一覧

MITI コード	日本語名称	実験でのラベル
GI	情報提供	0
Persuade	説得	1
Persuade with	許可を得た説得	2
Q	質問	3
CR	単純な聞き返し	4
SR	複雑な聞き返し	5
AF	是認	6
Seek	連携関係の探索	7
Emphasize	自立性の強調	8
Confront	直面化	9
NC	コード無し	10

本研究と[3]の異なる点は、ラベルなし英語面接記録を用いた半教師あり学習への拡張、英語のテキストデータの使用、面接者のみの評価、Bert を軽量化した Albert の使用、この 4 点である。

3 提案手法

Albert を用いることによって、分類タスクを行う。また、使用可能なデータセットを増やすために、教師なしデータを用いる半教師あり学習を用いる。具体的にはまずラベル付きデータにより学習を行う。次に得られたモデルを用い、ラベルのないサンプルへ予測結果をラベルとして設定する。これらのデータセットを統合して学習することで予測精度の向上を目指す。ラベルのついていないデータを用いる数とそのデータに対し擬似的なラベルを付与する回数を試行し、結果を考察する。今回はラベルのついていないデータの数を 10, 50, 100, 150 までで変化させ、ラベルを付与する回数を 1 から 5 までとした。

Consideration to Improve Prediction Accuracy of MITI Codes in Motivational Interviewing Using Semi-Supervised Learning
Sho UEMA, Computer Science and Intelligent Systems Program, University of the Ryukyus
Naruaki TOMA, Computer Science and Intelligent Systems Program, University of the Ryukyus

4 実験

4.1 実験手順

- 手順 1: データセットの準備。サイト[4]からラベル付けデータ、サイト[5]からラベル付けがされていないデータセットをダウンロード。
- 手順 2: 学習済みモデルのダウンロード('albert-base-v2')
- 手順 3: 交差検証を用いて学習。すでに作成された学習済みデータでテキストデータをベクトルに変換し、学習させる。交差検証の割合は、80%を学習用データとし、残りをバリデーションデータとし、エポック数を 20 にした。
- 手順 4: 手順 3 で学習したモデルを用いて、ラベルがないデータに対し、ラベルを予測する。この予測したラベルデータを擬似的なラベルとし、新たな学習用データとする。
- 手順 5: 手順 3 で用いた学習用データと、手順 4 で新たに作成した学習用データを結合したデータで再び学習を行う。
- 手順 6: 手順 3 でテストしたデータと同じデータでテストし、半教師あり学習で性能が向上したかを確かめる。

4.2 結果

いくつかの条件下で精度が向上したが、最も精度が向上した条件は追加するデータ数は 10、ラベル付けの回数は 3 回の時であった。半教師あり学習を行う前と後での混同行列を図 1, 2 に示す。

	0	1	2	3	4	5	6	7	8	9	10
0	0	0	0	1	0	2	0	0	0	0	0
1	0	0	0	0	0	2	0	0	0	0	0
2	0	0	0	0	0	1	0	0	0	0	1
3	0	0	0	8	0	1	0	0	0	0	2
4	0	0	0	0	3	10	0	0	0	0	0
5	0	0	0	2	4	19	0	0	0	0	0
6	0	0	0	1	0	1	0	0	0	0	0
7	0	0	0	1	0	0	0	0	0	0	1
8	0	0	0	1	0	0	0	0	0	0	0
9	0	0	0	2	0	0	0	0	0	0	0
10	0	0	0	0	0	3	0	0	0	0	13

図 1 擬似ラベル追加前の混同行列

4.3 考察

1 回の実験ごとに 5 回の擬似ラベルの付与を行っている。10 個の擬似ラベルを付与した場合は 5 回の実験中 4 回、50 個では 1 回、100 個では 5 回、150 個では 4 回、予測精度の改善に成功した。擬似ラベルを 100 個付与した場合は、精度が改善した回数が 5 回と最も多い。これより、擬似ラベルの数を増やすことが、必ずしも精度の改善に繋がるとは言えな

いということがわかる。

また、これらの擬似ラベルを付与するデータは先頭から順に抽出している。そのため、例えば 50 個のデータのうち 10 個のデータは重複がある。仮に精度を改善するようなデータが存在するとするならば、一度改善した実験で用いたデータセットを含むようなデータセットであるならば、それを含む様なデータセットであれば改善するはずである。しかし、その様な傾向は見られなかった。このことから、精度改善に寄与するためにはデータセットの適切な組み合わせと、不適切なデータの排除が必要だと考えられる。

	0	1	2	3	4	5	6	7	8	9	10
0	3	0	0	0	0	0	0	0	0	0	0
1	0	2	0	0	0	0	0	0	0	0	0
2	0	0	2	0	0	0	0	0	0	0	0
3	0	0	0	11	0	0	0	0	0	0	0
4	0	0	0	0	13	0	0	0	0	0	0
5	0	0	0	1	0	24	0	0	0	0	0
6	1	0	0	0	0	0	0	0	0	0	1
7	0	0	0	0	2	0	0	0	0	0	0
8	0	0	0	1	0	0	0	0	0	0	0
9	0	0	0	2	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	16

図 2 擬似ラベル追加後の混同行列

5 今後の展望

図 2 を見ると、半教師あり学習を用いたことにより、0, 1, 2 のラベルについて正しく予測できている。しかし、6, 7, 8, 9 のラベルについては正しく予測できたものが一つもない。またラベルの割合という観点で見ると、これらのラベルの割合は 1%程度しか変わらない。これらのラベルが予測できた理由及びできなかった理由を検討していきたい。

参考文献

- [1] ユースアドバイザー養成プログラム(改訂版), https://www8.cao.go.jp/youth/kenkyu/h19-2/html/ua_mkj.html, 第 5 章 5 節, 2021 年 10 月 29 日
- [2] T.B. Moyers, J.K. Manuel, D. Ernst, Motivational Interviewing Treatment Integrity Coding Manual 4.2.1, 2015 年 6 月
- [3] 津波 大輝, 動機付け面接における面接者支援システムの提案, 琉球大学 工学部 工学科 知能情報コース 令和 2 年度 卒業論文
- [4] Motivational Interviewing Treatment Integrity (MITI) code 3.0, <https://casaa.unm.edu/code/miti.html>, 2021 年 10 月 29 日
- [5] Downloads / Demos - Language and Information..., <https://lit.eecs.umich.edu/downloads.html#undefined>, 2021 年 10 月 29 日