

ファインチューニング手法 **ReFT** を用いた「情報」入試問題に対する 正答率向上の可能性検討

日本大学 文理学部

情報科学科 谷研究室

5421053 高野拓楽 54200 居関拓也

目次

1.はじめに

- 1.1 研究背景・動機
- 1.2 演習目的

2.準備

- 2.1 ReFT
- 2.2 学習データ
- 2.3 使用モデル
- 2.4 学習方法

3.演習

- 3.1 演習環境
- 3.2 演習結果

4.まとめ

- 4.1 考察と今後の課題

目次

1.はじめに

- 1.1 研究背景・動機
- 1.2 演習目的

2.準備

- 2.1 ReFT
- 2.2 学習データ
- 2.3 使用モデル
- 2.4 学習方法

3.演習

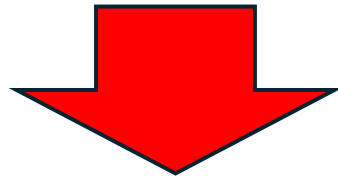
- 3.1 演習環境
- 3.2 演習結果

4.まとめ

- 4.1 考察と今後の課題

1.1 研究背景・動機

Chat-GPTなど一部のモデルは「何を」「どうやって」「どのくらい」学習したかが不明



自分たちで学習させて明確にする

1.1 研究背景・動機

- LLMを使用して大学入学共通テストの問題を解かせる試みがあった
- 自分たちもやってみよう
ただし使うモデルはオープンソースでファインチューニングできるものでなければならない

1.2 演習目的

共通テストや模試で「情報1」を学習させ、ファインチューニング後のモデル性能を検証

目次

1.はじめに

- 1.1 研究背景・動機
- 1.2 演習目的

2.準備

- 2.1 ReFT
- 2.2 学習データ
- 2.3 使用モデル
- 2.4 学習方法

3.演習

- 3.1 演習環境
- 3.2 演習結果

4.まとめ

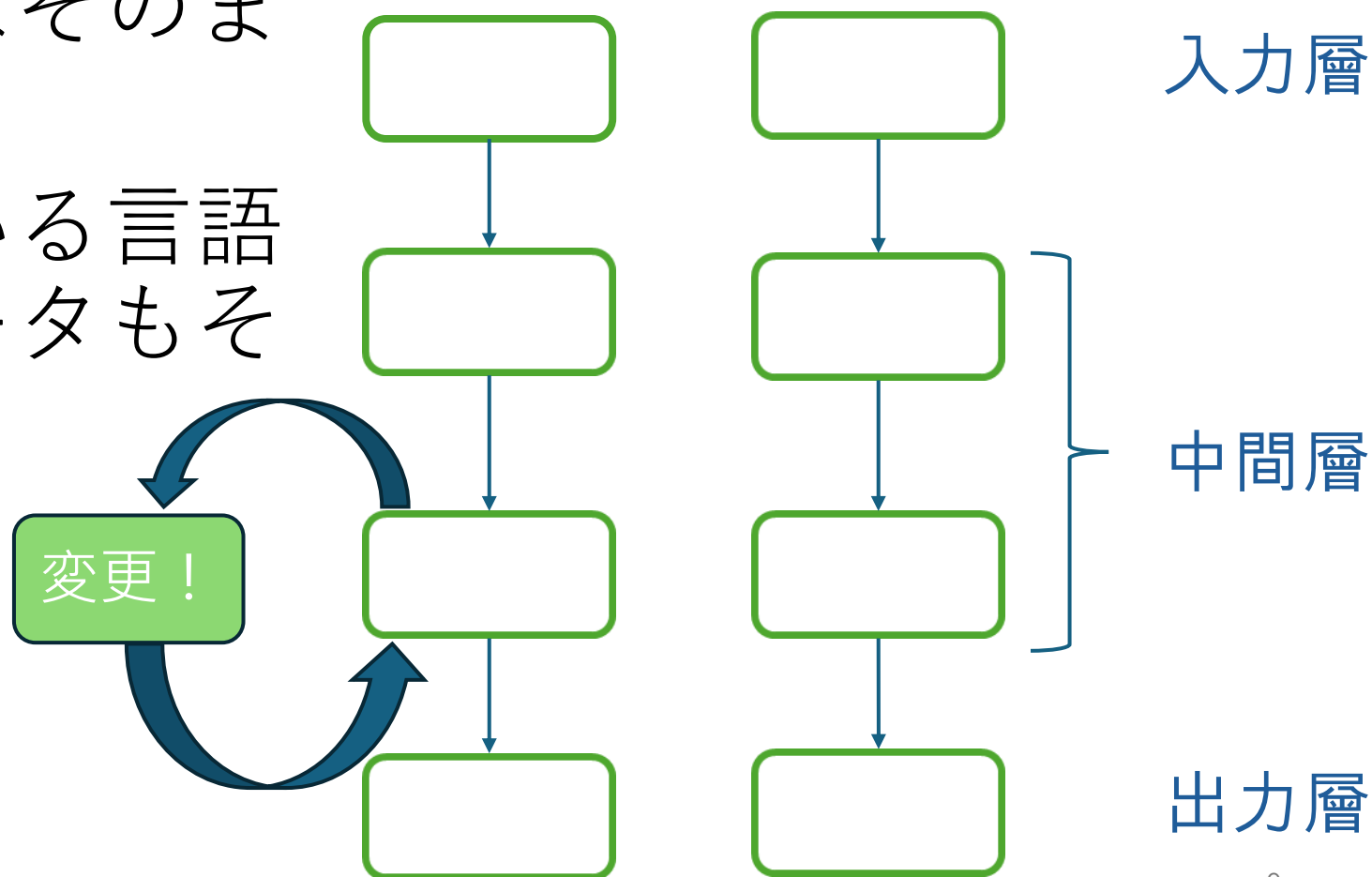
- 4.1 考察と今後の課題

2.1 ReFT

- スタンフォード大学のZhengxuan Wu（呉正玄）, Aryaman Arora（アリヤマン・アローラ）ら7名
- 2024年4月4日に発表
- LoRAよりも15～65倍効率的

2.1 ReFT

- 入力層と出力層はそのまま
- ベースになっている言語モデルのパラメータもそのまま
- 中間層に介入



2.2 学習データ

- EMIU
- 情報入試試作問題
- 2025年共通テストの情報 1

この中から「コンピュータとプログラミング」の分野のものを抽出

2.2 学習データ

EMIU

- 初等中等教育における情報分野の学習効果の評価手法の確立のために実施された，マークシート方式の情報模試
- 今回の実験では，P問題を使用
 - P問題が「コンピュータとプログラミング」の分野

2.2 学習データ

令和 7 年度大学入学共通テスト試作問題 情報 1

- 令和 7 年度入試から導入される「情報 1」の具体的なイメージ共有のために公開されたもの
- 使用した問題は大問 3 のみ
 - 大問 3 が「コンピュータとプログラミング」の分野

2.2 学習データ

令和 7 年度大学入学共通テスト 情報 1

- 使用した問題は大問 3 のみ
 - 大問 3 が「コンピュータとプログラミング」の分野

2.2 学習データ

- CSV形式
- 空欄がそのままのものと、空欄を正解で埋めたものがセット

Ex)

prompt	completion
“その際にいくら残るかは【力】で求めることができますね。”	“その際にいくら残るかは【46 % 10】で求めることができますね。”

2.3 使用モデル

rinna/llama-3-youko-8b

- 開発元：rinna
 - 対話型AIアバター、音声合成、LLM開発、アバター動画制作を行っている
- Llama-3-8Bに対して日本語データで継続事前学習を行ったモデルで、2024年5月7日に公開された

<https://huggingface.co/rinna/llama-3-youko-8b>

2.3 使用モデル

rinna/llama-3-youko-8b

- 事前学習のデータ
 - Japanese CC-100
 - Japanese C4
 - Japanese OSCAR
 - The Pile
 - Wikipedia
 - rinnaが精選した日本語のデータセット

モデルの選定

- 実験に使用した環境で動作すること
- 日本語に対応していること

モデルの選定

- 実験に使用した環境で動作すること
 - 実験環境で動作しない場合実験ができないため
 - Ex) rinna/llama-3-youko-70b-instruct はパラメータが多く、動作しなかった
- 日本語に対応していること

モデルの選定

- 実験に使用した環境で動作すること
 - 実験環境で動作しない場合実験ができないため
 - Ex) rinna/llama-3-youko-70b-instruct はパラメータが多く、動作しなかった
- 日本語の事前学習をしていること
 - 学習データや生成したいものが日本語のため

2.4 学習方法

交差検証

- データセットを4つに分割し3つの訓練データと1つのテストデータとする
- 3つの訓練データで学習させ残りのデータを解かせる
- この操作を学習させるデータを変えて4回繰り返す

2.4 学習方法

- 学習を始める前に事前学習として，答え方を学習させる目的でCSVファイルを読み込ませた
- 形式は，実際に学習に使用したデータと同じ
- 簡潔で短い問題が複数用意されている

prompt	completion
“その際にいくら残るかは【力】で求めることができますね。”	“その際にいくら残るかは【46 % 10】で求めることができますね。”

目次

1.はじめに

- 1.1 研究背景・動機
- 1.2 演習目的

2.準備

- 2.1 ReFT
- 2.2 学習データ
- 2.3 使用モデル
- 2.4 学習方法

3.演習

- 3.1 演習環境
- 3.2 演習結果

4.まとめ

- 4.1 考察と今後の課題

3.1 演習環境

- OS : Windows11
- GPU : NVIDIA GeForce RTX 4070 Laptop GPU
- CPU : 13th gen Intel(R) Core(TM) i7-13620H
- ストレージ :
- Python : 3.10.12

3.2 演習結果

学習前のモデルの解答

programmologyプログラムを組み立てましょう！
programmology プログラム解法を学ぶ(本) 創刊記念セール！ 40%OFF! 2020/6/30まで。
創刊記念！ 35%OFF! / programmology プログラム解法を学ぶ 第3版
namzn.to2,640円(2021年03月08日 21:23時点
Amazon調べ-詳細)(2021-03-08)"

学習前

- 選択肢だけではなく，問題も読み取れていない
- 「プログラム」という単語は認識している

3.2 演習結果

学習後のモデルの解答 一部抜粋

kakikokusekiを実行すると、以下のように表示
得られた結果に疑問があり、kakikokusekiを再
なった。(2)プログラムが正常に実行され、正
実行した。その後、一度も実行していなかった
表示されたが、出力結果に疑問があった。その
(1)と同じである。問4 次のプログラム(図
(8)、また、図2のプログラムの空欄①④⑤に
1(2)①: 9 ②: 4 ③: 1 ④: 99(3)③: 1
めには、どの空欄A~Dに入れるとよいか、解
最小交換確率の計算も行ったほうがよいと

ら、次のような表示が出た場合にSさんはうれしいと感じ
グラムが正しい表示をしたので、不具合もなくプログラム
確認しました。(3) 次の課程A~Eの課程を示せ。まず、
かからSさんの空欄をもとにして入れたらうれしいだろう
る。【ラ】、【リ】、【ロ】、【ハ】、【ニ】を選べ。た
B: 4 C: 2 D: 9 E: 1 【ラ】、【リ】、【ロ】、
【ハ】: ③④ 解答群、【ニ】: ③④ A: 9 回答(3.0 p
る。しかしそれは、発生に至る原因はSさんだけの責任で
らなかったし、Sさんのプログラムは実際にはプログラム

学習後

- 情報系の問題だという事は認識している
- 問題を解くという事は認識できていない

3.2 演習結果

学習後のモデルの解答

kakikokusekiを実行すると、以下のように表示して得た。(1)プログラムが正常に実行され、正しく表示されたのでいい結果が出力された。得られた結果に疑問があり、kakikokusekiを再度実行した。その後、一度も実行していなかったらしく、再度プログラムを表示するようになった。(2)プログラムが正常に実行され、正しく表示されたのでいい結果が出力された。得られた結果に疑問があり、kakikokusekiを再度実行した。その後、一度も実行していなかったらしく、再度プログラムを表示するようになった。(3)プログラムが正しく実行され、正しく表示されたが、出力結果に疑問があった。その後、プログラムを実行してみても表示が増えなかった。よってプログラムを表示した回数は(1)と同じである。問4 次のプログラム(図2)の実行結果にあった現象を、次の「解答群」のうちから選び、解答群の番号をすべて記し(8)、また、図2のプログラムの空欄①④⑤に対し、そのときの値を解答群から選び、解答群の番号をすべて記したい。(1) ①: 1 ②: 1 (2) ①: 9 ②: 4 ③: 1 ④: 99 (3) ③: 1 ④: 50 ⑤: 100 問5 次のプログラムで、問題があると疑問があがった。疑問を解決するためには、どの空欄A~Dに入れるとよいか、解答群のなかからすべてのものを選び、1)~6)の課程A~Dを示せ。また、疑問の解決には、最小交換硬貨枚数の計算も行ったほうがよいといわれて、疑問を解決するために入れるべきものも検討した。解答をすべて示せ。(2) 問6 疑問を検討した結果、疑問が解決し最小交換硬貨枚数も計算した。今のところ、Sさんはうれしいと感じている。いまのところ、は何かということだが、もし疑問が解決し、Sさんのプログラムが正しく表示している場合に限り、Sさんのうれしい気持はまだ続くはずであるから、次のような表示が出た場合にSさんはうれしいと感じると思われる。Sさんは、疑問の解決に時間をかけ、疑問を解決したあと、プログラムが正しい表示をしたので、不具合もなくプログラムが実行できたうれしさとともに、疑問が解決したうれしさでプログラムの表示を確認しました。(3) 次の課程A~Eの課程を示せ。まず、プログラムが作れたようだという発想で考えると、【ラ】には、【リ】の答のなかからSさんの空欄をもとにして入れたらうれしいだろう。プログラムの表示が確認されたときには、課程Eが表示されるはずである。【ラ】、【リ】、【ロ】、【ハ】、【ニ】を選べ。ただし、課程Eが表示されたら、Sさんはうれしいと思われる。「解答群」 A: 5 B: 4 C: 2 D: 9 E: 1 【ラ】、【リ】、【ロ】、【ハ】、【ニ】 解答群、【リ】: ② ④ ⑤ 解答群、【ロ】: ④ ⑤ 解答群、【ハ】: ③ ④ 解答群、【ニ】: ③ ④ A: 9 回答(3.0 pt/回答) 問7 「枚数(金額)」という関数には、実行結果にあった現象が発生する。しかしそれは、発生に至る原因はSさんだけの責任ではない。その原因については、Sさんにも原因があるが、その原因をSさんは知らなかったし、Sさんのプログラムは実際にはプログラムに不具合が出ていなかったからわからな

目次

1.はじめに

- 1.1 研究背景・動機
- 1.2 演習目的

2.準備

- 2.1 ReFT
- 2.2 学習データ
- 2.3 使用モデル
- 2.4 学習方法

3.演習

- 3.1 演習環境
- 3.2 演習結果

4.まとめ

- 4.1 考察と今後の課題

4.1 考察と今後の課題

考察

- 今回の演習で、評価指標を入れなかったため学習が適切におこなわれなかった可能性
- 事前学習に英語などが含まれていた為、プログラムのアルファベットが英単語として認識された可能性

今後の課題

- 評価指標を探し、導入
- 学習方法の改善