

2024年1月31日

## 日本語精度が高い 130 億パラメータの大規模言語モデル(LLM)を開発 ～ お客様の業種・業務に合わせたカスタム LLM を 2024 年春から提供開始 ～

株式会社リコー(社長執行役員:大山 晃)は、日本企業の業務での活用を目的に、企業ごとのカスタマイズを容易に行える 130 億パラメータの大規模言語モデル<sup>\*1</sup>(LLM)を開発しました。日本語と英語での学習において、その学習データの比率を工夫することで、日本語としての文法や回答が正確で日本語精度の高い、日本企業が持つ情報資産の活用に適したモデルを実現しました。ベンチマークツールを用いた性能検証<sup>\*2</sup>の結果、日本語で利用できる 130 億パラメータを持つ日本語 LLM において、2024 年 1 月 4 日現在で最も優れた結果を確認しました。

労働人口減少や高齢化を背景に、AI を活用した生産性向上や付加価値の高い働き方が企業成長の課題となっており、その課題解決の手段として、多くの企業が AI の業務活用に注目しています。しかし、AI を実際の業務に適用するためには、企業固有の用語や言い回しなどを含む大量のテキストデータを LLM に学習させ、その企業独自の AI モデル(カスタム LLM)を作成する必要があります。このたびリコーが開発した LLM は、米 Meta Platforms 社が提供する「LLM Llama2-13B」をベースに、日本語と英語のオープンコーパス<sup>\*3</sup>を追加学習させて開発したものです。①学習に利用するコーパスの選定 ②誤記や重複の修正などのデータクレンジング ③学習データの順序や割合を最適化するカリキュラム学習 など、リコー独自の学習上の工夫が組み込まれていることが特徴です。学習の結果、特に NLI(自然言語推論能力)において高性能となっています。日本語 LLM の性能評価で広く使われている日本語ベンチマークツール(llm-jp-eval)を用いた他 LLM モデルとの性能比較<sup>\*2</sup>では、評価スコアの平均値が最も高く、優れた性能を確認することができました。

モデル	Avg.	MC	NLI	QA	RC	Jam p	JaN LI	JCom monse nseQA	JEM Hop QA	JNL I	JSe M	JSI CK	JSQ uAD	NII LC
Llama 2 13B	58.9	67.7	39.6	43.2	85.1	41.1	54.3	67.7	44.4	15.8	55.9	30.7	85.1	42.0
Llama 2 13B Chat	57.9	62.5	47.3	37.9	84.0	<b>46.3</b>	54.6	62.5	43.8	31.4	65.7	38.4	84.0	32.0
PLaMo-13B	47.4	22.6	44.0	46.6	76.2	35.9	50.0	22.6	<b>51.9</b>	14.7	59.1	60.4	76.2	41.4
LLM-jp-13B	46.3	22.3	42.5	43.1	77.4	37.6	49.7	22.3	47.9	31.7	67.2	26.1	77.4	38.3
Stockmark-13B	-	22.2	-	-	73.0	-	49.6	22.2	-	-	-	-	73.0	50.2
Swallow-13B	66.2	<b>78.0</b>	39.7	<b>56.9</b>	<b>90.0</b>	36.2	50.0	<b>78.0</b>	50.5	34.0	49.7	28.4	<b>90.0</b>	63.3
ELYZA-japanese-Llama-2-13b-fast	62.5	74.3	44.9	45.0	86.0	35.9	52.9	74.3	45.3	34.1	<b>68.4</b>	32.9	86.0	44.6
Ricoh-13B (ours)	<b>67.0</b>	72.7	<b>50.4</b>	56.3	88.5	34.2	<b>56.9</b>	72.7	49.0	<b>34.2</b>	52.2	<b>74.6</b>	88.5	<b>63.6</b>

日本語ベンチマークツール(llm-jp-eval)における他モデルとの比較結果<sup>\*2</sup>(リコーは最下段)

株式会社リコー <https://jp.ricoh.com/>

報道関係のお問い合わせ先 広報室 TEL : 050-3814-2806 (直通) E-mail : [koho@ricoh.co.jp](mailto:koho@ricoh.co.jp)  
 お客様の問い合わせ先 仕事のAI お問合せフォーム  
[https://www.secure.rc-club.ricoh.co.jp/shigoto-no-ai\\_inq?](https://www.secure.rc-club.ricoh.co.jp/shigoto-no-ai_inq?)

学習能力が高い本 LLM に企業独自の情報や知識を取り入れることで、お客様ごとの業種・業務に合わせた高精度な AI モデル(カスタム LLM)を、短期間で容易に構築することが可能になります。カスタム LLM をお客様の業務で活用し、業務文書の要約や質問応答の作業を AI に置き換えることで、業務のワークフローを最適化し、業務効率化を実現します。

2024 年春から、カスタム LLM をクラウド環境でお客様へ提供開始します。まずは日本国内のお客様より提供を開始し、今後海外のお客様への提供も目指します。

LLM は、パラメータ数が多いほどより多くの情報を処理でき、学習能力が高まる一方、より大きな計算リソースが必要になるうえ、学習や運用の難易度が上がり、開発難易度が高くなります。本モデルは「AWS<sup>™</sup> LLM 開発支援プログラム」によるサポート提供のもと、AWS Trainium アクセラレーターを搭載した Amazon Elastic Compute Cloud Trn1 インスタンスを利用することで、従来手法の開発と比較し 45% のコスト低減および 12% の開発期間短縮を実現することができました。さらに、お客様向けカスタム LLM を開発する際にも、効率的に開発することができるため、お客様へより安価・短納期でのご提供が可能です。

なお、今回の LLM 開発とその手法は 2024 年 3 月 11 日から開催される言語処理学会\*4 第 30 回年次大会において論文発表する予定です。

リコーは、お客様に寄り添い、業種業務に合わせて利用できる AI サービスの提供により、お客様が取り組むオフィス／現場のデジタルトランスフォーメーション(DX)を支援してまいります。

\*1 人間が話したり書いたりする言葉(自然言語)に存在する曖昧性やゆらぎを、文章の中で離れた単語間の関係までを把握し「文脈」を考慮した処理を可能にしているのが特徴。「自然文の質問への回答」や「文書の要約」といった処理を人間並みの精度で実行でき、学習も容易にできる技術。

\*2 2024 年 1 月 4 日時点の評価結果。公開されている日本語 LLM のうち 130 億パラメータかつトークナイザを日本語適応している事前学習モデルとの比較。

MC、NLI、QA、RC は、下記カテゴリ内の平均値。Avg.はその 4 カテゴリ平均値の平均値。

MC (Multi-Choice QA: 多肢選択質問応答): jcommonsenseqa

NLI (Natural Language Inference: 自然言語推論): jamp, janli, jnli, jsem, jsick

QA (Question Answering: 質問応答): jemhopqa, niilc

RC (Reading Comprehension: 読解): jsquad

\*3 自然言語の文章や使い方を大規模に収集し、一般公開されているデータセット。

\*4 自然言語処理(NLP)分野では最大の学会です。学会発表する NLP2024 は、言語処理学会の年次大会です。

<https://www.anlp.jp/>

## ■ 関連情報

企業独自の AI モデルを簡単に作成できるノーコード開発ツールのトライアル提供を開始

[https://jp.ricoh.com/release/2023/0710\\_1](https://jp.ricoh.com/release/2023/0710_1)

※AWSは米国その他の諸国における、Amazon.com, Inc.またはその関連会社の商標です。

※社名、製品名は、各社の商標または登録商標です。

---

## ｜ リコーグループについて ｜

リコーグループは、お客様のDXを支援し、そのビジネスを成功に導くデジタルサービス、印刷および画像ソリューションなどを世界約200の国と地域で提供しています(2023年3月期グループ連結売上高2兆1,341億円)。

”はたらく”に歓びを 創業以来85年以上にわたり、お客様の“はたらく”に寄り添ってきた私たちは、これからもリーディングカンパニーとして、“はたらく”の未来を想像し、[ワークプレイスの変革](#)を通じて、人ならではの創造力の発揮を支え、さらには持続可能な社会の実現に貢献してまいります。

詳しい情報は、こちらをご覧ください。

<https://jp.ricoh.com/>