

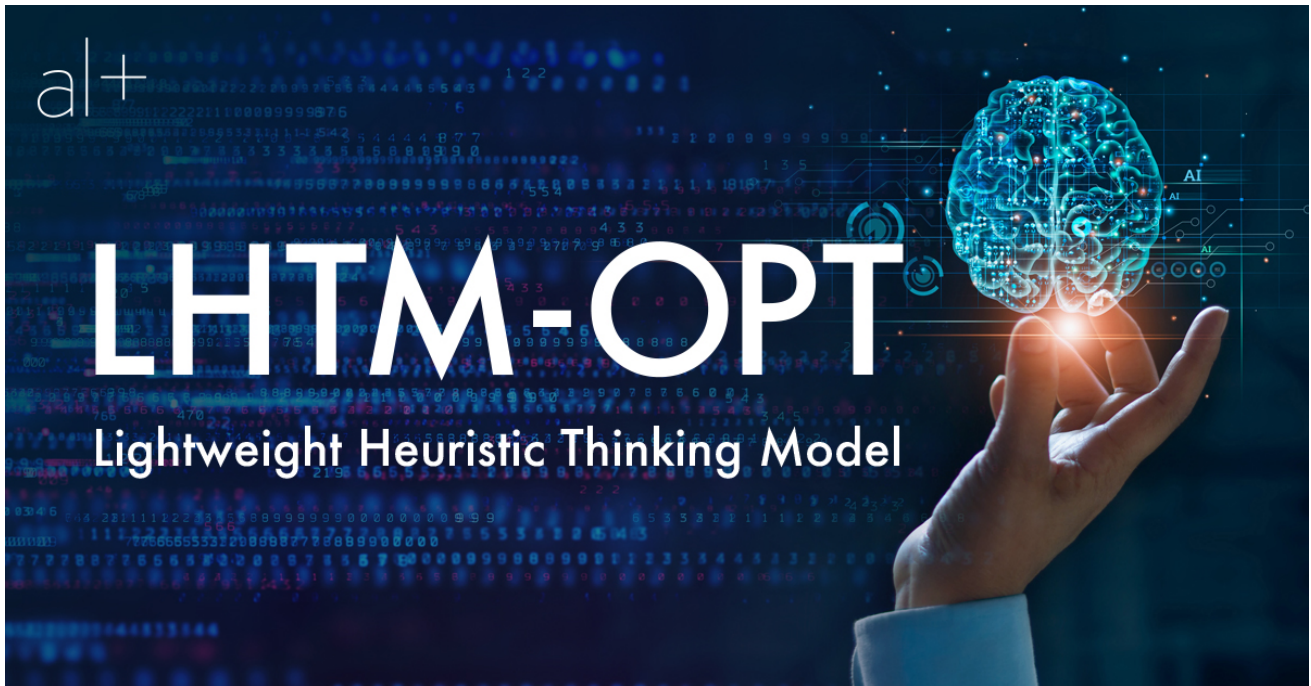
ニュースリリース

オルツ、軽量かつ高精度を実現した軽量型大規模言語モデル「LHTM-OPT」を発表 ～日本語LLMを評価する「Rakuda」ベンチマークで最高スコアを記録し、 国産商用プライベートLLMとして国内最高に～

P.A.I.®（パーソナル人工知能）をはじめ、AIクローン技術でつくり出すパーソナルAIの開発および実用化を行う株式会社オルツ（本社：東京都港区、代表取締役：米倉 千貴、以下、オルツ）は、小規模GPUマシンで実用的な、パラメータ数が最適化された新たな大規模言語モデル「LHTM-OPT（ラートム・オプト）」を開発し、日本語商用プライベートLLM（大規模言語モデル）として最高性能を記録したことを発表いたします。

※国内最高・最高性能：

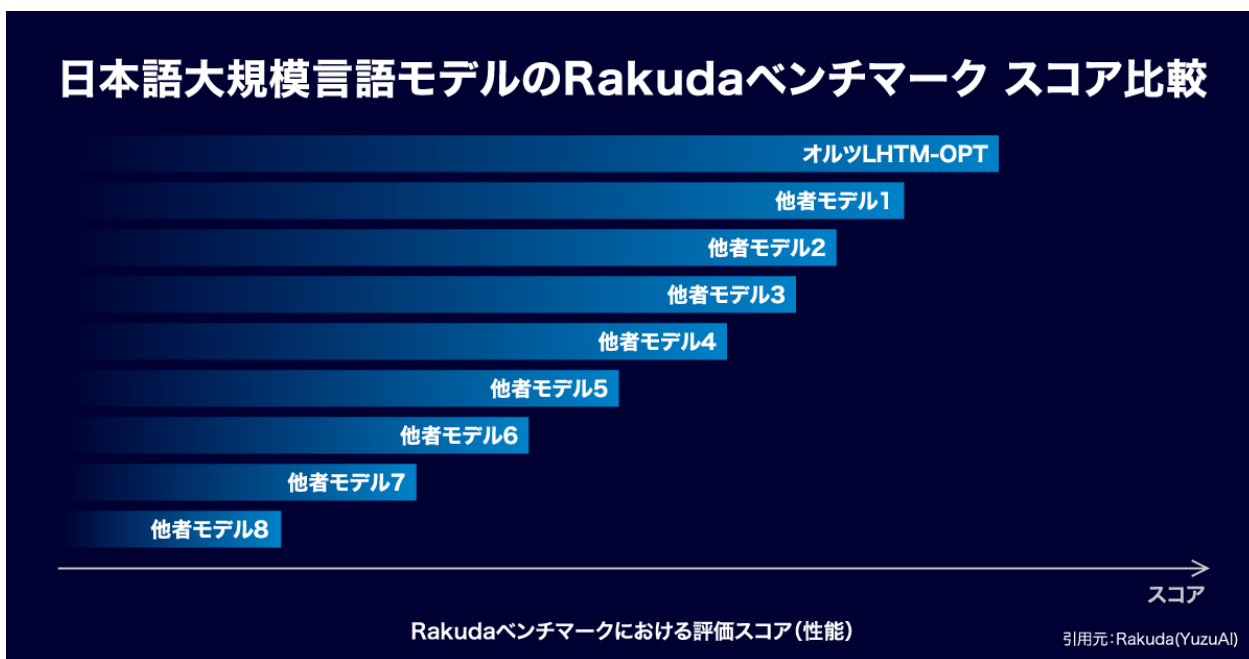
「Rakuda」LLMベンチマークによる評価で、国内の大規模言語モデルの中での比較として国内トップ性能を達成。
(2023年10月19日時点。自社調べ)



大規模パラメータの大規模言語モデルは膨大な計算リソースとエネルギーを必要とするため、環境に対する負荷の増大が懸念されています。そのため、現在は軽量でありながら大規模モデルと同等の精度を実現するための小規模パラメータモデルの研究が活発化しています。このアプローチは、大規模言語モデルから重要な知識を抽出して、新たに有用なモデルを開発することから「知識蒸留」と表現されており、計算リソースの削減やエネルギー効率の向上に寄与する試みです。オルツは、このような社会課題を解決する生成AI開発に積極的かついち早く取り組みはじめ、小規模パラメータモデルである「LHTM-OPT」の開発に成功しました。

当社が提供する大規模言語モデル「LHTM-2」 (<https://alt.ai/news/news-1892/>) は、パラメータ数が膨大で調整の自由度が非常に高く設定されています。それに対して、「LHTM-OPT」は、そのパラメータを小規模GPUマシンでも実行可能であり、実用に最適化された軽量型モデルです。

「LHTM-OPT」の開発にあたり、当社はオープンデータの学習だけでなく、当社独自の指示データや対話データのデータセットを作成し、精度を改善してまいりました。そして、YuzuAIグループが出した日本語LLMを評価するRakudaベンチマーク (<https://yuzuai.jp/benchmark>) では、「LHTM-OPT」の評価スコアが1152になり、2023年10月19日時点で日本国内で最高レベルに達しました。また、日本語言語理解ベンチマーク「JGLUE (Japanese General Language Understanding Evaluation)」8タスク平均においても、商用モデルとして国内最高スコア56.12を記録しました（「LHTM-OPT」のJGLUE 主要4タスク平均のスコアは75.08になっています）。10Bパラメータ未満においてそれぞれのスコアを実現できたことは、グローバルレベルにおける当社の技術力の高さを示しているものと考えております。今後も継続して様々なベンチマークで評価を行い、モデルをブラッシュアップすることで「LHTM-OPT」の性能と汎用性を向上させ続けるとともに、本モデルを基盤として自動質問応答、議事録自動要約、情報抽出、会話理解、予測分析のデータ整理・作成などの高度な処理が必要な問題に対応可能な多くのアプリケーションを実装していきます。



<軽量版大規模言語モデルが高精度を実現した意義とクライアントへの価値提供可能性>

- **意義：軽量かつ高精度を実現する国内最高レベルの「知識蒸留」技術**
 - 今回の「知識蒸留」を含めた学習と評価過程において、当社はデータ企画から作成、GPUリソースの評価・確保、学習プロセス最適化など一貫した企画・技術開発・運用を実践しています。その結果として、軽量かつ高精度なモデルを実現することで、より実践的な生成AIプロダクトの基盤構築に対するオールツ技術の高い品質を証明しています。
- **提供価値：カスタマイズ容易性**
 - 当社のクライアントは、今回開発した軽量かつ高精度なモデルを専門分野でファインチューニング可能です。金融や製薬、製造、不動産など様々な分野固有の情報をベースにカスタマイズすることで、パラメータ数が圧倒的に多いLLMと同等の回答精度を保持しながら専門性を持たせることが可能になります。
- **提供価値：GPUコストの節約**
 - 当社のクライアントは、大規模言語モデルを基盤としたサービスを構築する際に課題となるランニングコストについても最適化されます。軽量かつ高精度なモデルを運用することで、継続して活用する際のコスト対効果を最大化することが可能になります。

当社は、世界水準の技術で "アジアにおけるOpenAI" の地位確立を目指すとともに、顧客へ最高品質のソリューションを提供することで、日本企業の労働生産性向上に資する取組みを推進して参ります。

LLMを活用したプロジェクトについては、以下のアライアンス窓口にお問い合わせください。

▶LHTM-2/LHTM-OPT/GPT など大規模言語処理ソリューションに関するお問い合わせ先

<https://alt.ai/aiprojects/gpt/>

■株式会社オルツについて

2014年11月に設立されたオルツは、P.A.I.®（パーソナル人工知能）、AIクローンをつくり出すことによって「人の非生産的労働からの解放を目指す」ベンチャー企業です。また、AIの対話エンジンの開発から生まれた音声認識テクノロジーを活用した「[AI GIJIROKU](#)」などのSaaSプロダクトを開発・提供しています。2023年9月までの累計調達額は80億円超に達しています。

<https://alt.ai/>

<報道関係者からのお問い合わせ先>

株式会社オルツ 広報 西澤

e-mail : press@alt.ai

<アライアンスに関するお問い合わせ先>

株式会社オルツでは、IT・金融・建設・物流・メディア・製造・小売・サービス業など、ジャンルを問わずAIソリューションの提供および支援を行っております。

お気軽にお問い合わせください。

株式会社オルツ AI Solutions事業部 浅井

e-mail : gptsolutions@alt.ai