



Nomura Research Institute Group

NEWS RELEASE

2025 年 4 月 15 日

株式会社野村総合研究所

野村総合研究所、業界・タスク特化型 LLM の構築手法を独自に開発

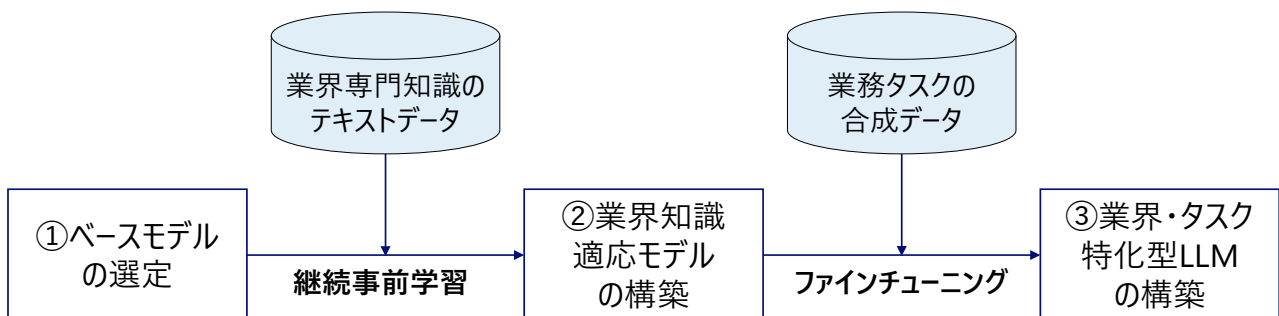
～小規模モデルで大規模商用モデルを超える精度を実現、多様な業界への展開が可能に～

株式会社野村総合研究所（本社：東京都千代田区、代表取締役 社長：柳澤花芽、以下「NRI」）は、独自の技術とノウハウを活用し、80 億パラメータの比較的小規模なモデルをベースに、特定業界やタスクに特化した大規模言語モデル（以下「業界・タスク特化型 LLM」）の構築手法を開発しました。本手法を用いて開発したモデルは、特定のタスク¹において、大規模な商用汎用モデルである GPT-4o を超える性能を示しています。また、本手法は多様な業界やタスクへの応用が可能です。

■ 業界・タスク特化型 LLM の構築手法の開発背景と成果

GPT-4o をはじめとする一般的な汎用モデルは幅広いタスクで利用可能な一方で、専門性の高いタスクや特定の業界で求められる高度な専門知識や独自の用語・規制などへの対応が難しく、加えてパラメータ数が大きく計算コストも高いという課題があります。NRI では、これらの課題を解決するため、低コストかつ高精度で実務に即した業界・タスク特化型 LLM の構築手法を開発しました。具体的には、以下の 3 段階のアプローチを採用しています。

図：業界・タスク特化型 LLM 構築の流れ



① 日本語性能に優れ、コスト効率の高い小規模ベースモデルの選定

日本語性能に優れた東京科学大学と産業技術総合研究所の「Llama 3.1 Swallow 8B」(80 億パラメータ)をベースモデルとして利用しました。「Llama 3.1 Swallow 8B」は小規模モデルであるため、計算リソースや運用コストを削減できます。

また、本手法では、オープンウェイト²の LLM を基盤としたことで、ベースモデルを特定のモデルに固定せずに、目的やタスクに応じて適切に選定することが可能です。将来的なモデルアップデートにも柔軟に対応できます。

② 継続事前学習による業界知識適応モデルの構築

銀行・保険等の金融業界を例に、日本語の専門知識テキストデータを日本語金融コーパスとして独自に構築し、継続事前学習³を実施しました。これにより、ベースモデルが持つ一般的な言語能力や知識を維持しつつ、業界特有の専門的知識を効果的に習得させる汎用的な仕組みを構築しました。

③ 合成データを用いたタスク特化型 LLM の構築

今回ターゲットとした保険業界の営業コンプライアンスチェックにおいては、規制違反を含む会話データの収集が困難であるため、LLM を用いて多様なシナリオを想定した合成データ⁴を生成しました。この合成データをもとにファインチューニング⁵を実施することで、該当タスクに特化した LLM を構築しました。

こうした取り組みにより、保険業界の営業コンプライアンスチェック試験では、商用大規模モデルの GPT-4o (2024-11-20) を 9.6 ポイント上回る正解率⁶を実現しています。

表：保険業界の営業コンプライアンスチェックの性能評価結果

モデル	正解率
GPT-4o (2024-11-20)	76.7%
ベースモデル (Llama 3.1 Swallow 8B Instruct v0.2)	51.7%
NRI 独自特化型モデル (Llama 3.1 Swallow 8B + ファインチューニング)	83.1%
NRI 独自特化型モデル (Llama 3.1 Swallow 8B + 継続事前学習 + ファインチューニング)	86.3%

また、今回の構築手法はプロセスとして標準化しており、他の業界やタスクにも適用が可能です。使用するデータや合成データの生成内容を変えることで、適用範囲を広げることができます。

■ 今後の展開と研究連携

NRI は今回の成果をもとに、他の業界やタスクへの最適化をさらに加速させる予定です。今回開発した業界・タスク特化型 LLM の特長である、小規模モデルによる高精度・低コスト・迅速で柔軟な適応性を活かし、汎用モデルでは対応が難しい専門領域への展開も推進していきます。

2025 年度には東京科学大学 岡崎研究室との共同研究を予定しており、具体的な業界課題を反映した実証実験やモデル技術の改良を重ね、生成 AI の社会実装を進めます。さらに、ビッグテック企業やスタートアップとの連携も強化し、技術の商用展開や実用化を推進することで、多様な業界や用途（タスク特化した AI エージェントへの活用等）での利活用を拡大していきます。

東京科学大学 情報理工学院

教授 岡崎 直観 氏からのコメント

東京科学大学と産業技術総合研究所では、日本語に強く、商用利用も可能な汎用型大規模言語モデルを目指し、Swallow シリーズを開発・公開してきました。このたび、業界・タスク特化型 LLM のベースモデルとして Llama 3.1 Swallow が利用され、独自の工夫により高い正解率を達成されたことを喜ばしく思っています。共同研究を通じて、業務特化型のモデルの開発手法をさらに洗練させ、社会実装を進めていきたいと考えています。

-
- ¹ 保険業界の営業コンプライアンスチェック（保険会社の営業担当者が法律や規則に従って適切に業務を行っているかを営業会話の履歴から確認すること）のタスクで試験を実施。
 - ² オープンウェイト (Open Weight): モデルのパラメータが公開されていること。
 - ³ 継続事前学習 (Continued Pre-Training): 既に一般的な知識を持つ事前学習済みモデルに対し、新たな領域やタスクに合わせた追加データでさらに学習させ、専門性や適応性を向上させる手法。
 - ⁴ 合成データ (Synthetic Data): モデルの学習に利用するために、LLM などの生成 AI で人工的に生成したデータ。今回は、Llama-3.1-Swallow-70B-Instruct-v0.3 を利用して生成。
 - ⁵ ファインチューニング (Fine-Tuning): 既に事前学習済みのモデルのパラメータを特定タスク向けに微調整し、タスク固有の性能を最適化する追加学習手法。
 - ⁶ 評価データ 344 件（7 種類の違反項目に抵触する会話 172 件と、いずれの違反項目にも抵触しない会話 172 件）に対して、違反項目（なしも含む）の正解率で評価を実施。

【ニュースリリースに関するお問い合わせ】

株式会社野村総合研究所 コーポレートコミュニケーション部 梅澤、弓野
TEL : 03-5877-7100 E-mail : kouhou@nri.co.jp

【本件に関するお問い合わせ】

株式会社野村総合研究所 AI ソリューション推進部 岡田、大河内
E-mail : ai-tech-lab@nri.co.jp