

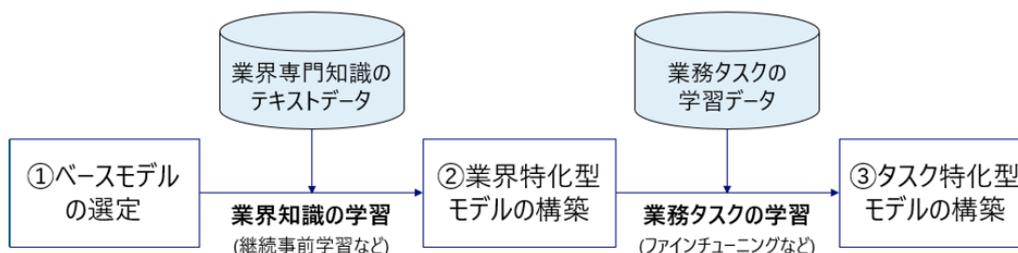
### 野村総合研究所、AI エージェント活用へ向けて 業界・タスク特化型 LLM の構築手法を高精度化・効率化

～GENIAC 第3期成果として、金融実務で商用大規模モデルを上回る精度を確認～

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、経済産業省と NEDO<sup>1</sup>が実施する国内の生成 AI の開発力強化を目的としたプロジェクト「GENIAC（Generative AI Accelerator Challenge）」第3期（2025年8月～2026年3月）の支援<sup>2</sup>を受け、特定の業界や業務に特化した大規模言語モデル（LLM）の構築手法を高精度化・効率化させました。

この LLM を AI エージェントの仕組みに組み込み金融業界の複数の業務で検証した結果、いずれも OpenAI 社が開発した ChatGPT 搭載の商用大規模モデル「GPT-5.2」を上回る精度を示しました。NRI はこの成果を活用し、多様な業界の企業に対し、AI エージェントを活用した専門業務支援に取り組んでいきます。

（図）業界・タスク特化型 LLM 構築の流れ



#### ■ GENIAC 第3期の研究開発背景と成果

汎用的な生成 AI モデルは幅広い業務で利用できる一方、特定の業界で求められる専門知識や独自の用語・規制への対応が難しいという課題があります。企業の実務では、法律やガイドライン等を踏まえた専門的な判断が求められる場面が多く、こうした業務を AI で支援するには、業界や業務ごとに最適化されたモデルや AI エージェントが必要です。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対応できる「業界・タスク特化型 LLM」の構築手法を独自に研究開発してきました。今回の GENIAC 第3期では、これまでの研究成果をさらに発展させ、以下の特徴をもつ構築手法を開発しました。

### ① ベースモデルの選定 ～基礎性能が高いベースモデルを選定・検証～

本手法では、一般に公開されている LLM を学習の土台（ベースモデル）として採用しています。これにより、特定のモデルに固定されず、目的や業務に応じて最適なモデルを選択でき、将来のモデル更新にも柔軟に対応できます。今回は基礎性能が高い中規模<sup>3</sup>の公開モデルを複数選定し、追加学習によってより専門的な知識を持ちながら、業務プロセスを高精度に処理できるかを検証しました。さらに、モデルの圧縮<sup>4</sup>によって少ない計算資源でも稼働でき、圧縮後も精度が大きくは低下しないことを確認しました。

### ② 業界特化型モデルの構築 ～学習データ収集および作成を自動化～

特定業界の専門知識をベースモデルに学ばせるため、二段階の仕組みを構築しました。第一に、業界知識を含む大量のテキストデータを自動的に収集・選別し、モデルに専門知識を習得させる仕組み<sup>5</sup>です。第二に、業務に即した問答データ<sup>6</sup>を高性能な LLM<sup>7</sup>を用いて自動生成し、モデルの応答精度を高める仕組み<sup>8</sup>です。これにより、対象とする業界を問わず、専門知識を備え、かつ質問に対して的確に回答できる業界特化型モデルを容易に構築できるようになりました。今回は金融業界を題材としてモデルを構築し、金融知識の理解度を測る標準テスト<sup>9</sup>でベースモデルを上回る精度を確認しました。

### ③ タスク特化型モデルの構築 ～学習データ作成を自動化～

特定の業務に特化させるための学習用データについても、LLM<sup>10</sup>を用いて自動生成する仕組みを構築しました。少数の業務サンプルと業務の定義があれば、対象とする業務を問わず学習用データを容易に生成できます。今回は証券・保険の複数の実務<sup>11</sup>を題材として検証した結果、いずれの業務においても NRI が構築した特化型モデルが商用大規模モデル GPT-5.2 を上回りました。

(表) 証券・保険分野の複数のタスクの性能評価結果（正解率）

モデル	証券営業会話チェック	保険営業会話チェック	保険募集文書校正
NRI 独自タスク特化型モデル	88.00%	94.12%	86.67%
ベースモデル (gpt-oss-20b)	78.44%	87.73%	70.48%
GPT-5.2 (参考)	85.56%	92.67%	66.67%

### ④ AI が自律的に業務を遂行する環境でも有効性を確認

モデルの検証にあたっては、AI が複数の手順を自律的に実行する「AI エージェント」の仕組みの中に特化型モデルを組み込み、実際の業務手順を模した環境で試験を行いました。その結果、業務プロセスの一部として AI が担当者の判断を支援する使い方にも対応できることを確認しました。

## ■ 研究開発成果の公開と今後の展望

NRI では、本研究成果の一部を、2026 年 3 月に開催された言語処理学会第 32 回年次大会（NLP2026）にて発表しました<sup>12</sup>。また、構築した業界特化型モデルと学習用データを Hugging Face<sup>13</sup>で公開しています。NRI は今後、企業の専門的な業務プロセスを AI で支援するため、業界・タスク特化型モデルを組み込んだ AI エージェントの導入・展開を進めていきます。本手法は金融以外の業界にも適用可能であり、多様な業界や業務への拡大を図ることで、汎用モデルでは対応が難しい専門領域の AI 活用を推進していきます。

---

<sup>1</sup> NEDO：国立研究開発法人 新エネルギー・産業技術総合開発機構。

<sup>2</sup> GENIAC（Generative AI Accelerator Challenge）：経済産業省と NEDO が 2024 年 2 月に開始した、国内の生成 AI 開発力強化を目的としたプロジェクト。生成 AI の基盤モデル開発において、日本が国際競争力を確保し、国内産業への実装を加速させることを目指しており、採択事業者に対して大規模な計算資源の提供支援等を行っている。第 3 期（2025 年 8 月～2026 年 3 月）は領域特化型のモデルなど、社会実装に繋がるテーマが重視されており、NRI が提案した「業界・タスク特化型 LLM 構築手法の開発」テーマも事業性が評価され採択された。

<sup>3</sup> 中規模モデル：100 億～400 億パラメータ規模のモデル。パラメータはモデルの学習によって調整される数値で、一般にパラメータ数が多いほど高い性能が期待できるが、運用コストも増大する。

<sup>4</sup> モデルの圧縮：ここでは量子化と呼ばれる技術を指す。モデル内部の数値の精度を下げることでデータ量を削減し、少ない計算資源で動作させる手法。

<sup>5</sup> 技術的には「継続事前学習（Continued Pre-Training）」と呼ばれる手法。既に一般的な知識を持つ学習済みモデルに対し、新たな領域に合わせた追加データでさらに学習させ、専門性を向上させる。

<sup>6</sup> 技術的には「指示データ（Instruction Data）」と呼ばれる学習用データ。指示文と、それに対する模範的な応答文の組み合わせで構成される。本研究では、応答文に思考の過程（Chain-of-Thought）を含めることで、モデルが複雑な判断を行う際に論理的な推論を経て回答する能力の向上を図っている。

<sup>7</sup> 問答データの自動生成には、学習元のベースモデルよりも高性能なオープンウェイトモデル（openai/gpt-oss-120b）を利用することで、高性能なモデルの能力を中規模モデルに再学習（蒸留）させている。

<sup>8</sup> 技術的には「指示チューニング（Instruction Tuning）」と呼ばれる手法。学習済みモデルに対し、指示と応答の組み合わせからなるデータで追加学習を行い、指示に沿った正確な応答を生成する能力を高める。

<sup>9</sup> japanese-lm-fin-harness（言語モデル性能評価のための日本語金融ベンチマーク。<https://github.com/pfnet-research/japanese-lm-fin-harness>）および pfmt-bench-fin-ja（金融分野に特化した複数ターン日本語生成ベンチマーク。<https://github.com/pfnet-research/pfmt-bench-fin-ja>）を対象に計測。

<sup>10</sup> 特定業務の学習用データの生成についても、同様に高性能なオープンウェイトモデル（openai/gpt-oss-120b）を利用。

<sup>11</sup> 証券・保険の営業会話チェック（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記録から確認すること）と、保険の募集文書校正（保険商品の募集文書が法律や規則に従って適切な記載や表現で作成されているかを確認すること）のタスクで評価を実施。評価データ数はそれぞれ 90 件（証券営業会話チェック）、344 件（保険営業会話チェック）、21 件（保険文書校正）で、5 回ずつ推論させた正解数の平均値（average@5）で評価。

<sup>12</sup> 言語処理学会（NLP2026）発表論文「ドメイン特化 LLM の推論能力向上を目的とした合成指示データセットの構築と金融ドメインにおける評価」（[https://www.anlp.jp/proceedings/annual\\_meeting/2026/pdf\\_dir/C7-2.pdf](https://www.anlp.jp/proceedings/annual_meeting/2026/pdf_dir/C7-2.pdf)）にて研究成果を発表。

<sup>13</sup> <https://huggingface.co/nri-ai> にて業界特化型モデルと学習用データ（問答データ）を公開。

**【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 瀬戸、弓野

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

**【ご参考】**

- GENIAC (METI/経済産業省) [https://www.meti.go.jp/policy/mono\\_info\\_service/geniac/index.html](https://www.meti.go.jp/policy/mono_info_service/geniac/index.html)
- 言語処理学会第 32 回年次大会 (NLP2026) <https://www.anlp.jp/nlp2026/>