

食品・流通小売領域の専門知識を取り込んだマルチモーダルな大規模モデル開発

実施者	NABLAS株式会社
事業概要	- 日本特有かつ流行をおさえた食品に強い生成 AI 技術の開発 - 社会実装に向け、サービス化するとともに、一部モデル・ノウハウ等の成果を公開 食品の関するトレンド分析や売り上げ・いいね数を用いた分析など、生成AIを用いた定性・定量分析を可能にするサービスの開発。 ▶ 日本の食品業界特有の知識を高度に扱える大規模視覚言語モデル (Large Vision-Language Model: LVLM) - 事前学習を行い、パラメーター数の異なるLVLM数種を開発 - 指示学習データセットの異なるLVLMを開発 - 汎用型 - 領域特化型
社会実装イメージ	以下の仕組みを「ツール」としてAIEージェントに提供し、エンドユーザーの入力に柔軟に対応できるUIを備えたサービスを展開。 ①SNSや関連サービスからデータを収集 ②開発した大規模モデルを用いて指定の形式に変換 ③データベースにデータを自動で蓄積する仕組みを構築

事業成果	汎用的な大規模視覚言語モデルNABLA-VLの開発および公開。 NABLA-VLは日本語と英語の両ベンチマークで優れたスコアを達成し、特に、図表を含んだ質問応答ベンチマークであるJDocQAでは29.15ptを達成してgpt-4o-2024-11-20とすべての国産モデルを上回った（5月19日時点）。 また、本事業中に作成した食に関するデータセット上でNABLA-VLの微調整を行い、NABLA-VL.foodの開発を行った。 NABLA-VL.foodを利用して食に関する口コミ収集や商品紹介を収集できるパイプラインを構築した。 さらに、このパイプラインを組み込み、食品会社や小売事業者の業務を効率化するツールを作成した。具体的な利用方法としては、トレンド抽出や口コミ抽出に加えて政府系のニュースの要約などが挙げられる。これらに加え、NABLA-VLの再現および微調整を行えるコードベースを公開した。 さらに、評価ツールやVisionZipを使ったトークン圧縮化手法（実験環境では学習時間を約2倍、推論時間を約1.3倍短縮）をこのコードベースに組み込んだ。また、Wikipediaの日本語ページに関する画像を含んだ質問応答データセットWikipedia-VQA-Jaを公開した。
------	---