

2025年03月21日
株式会社 NTT ドコモ

大規模言語モデルから生成したユーザーモデルを用いて実ユーザーの行動を予測する 仮想マーケティング技術を開発

～さまざまな観点で柔軟にシミュレーション可能な高効率マーケティングをめざす～

株式会社 NTT ドコモ（以下、ドコモ）は、企業固有のデータと大規模言語モデル（以下、LLM）により生成したユーザーモデルへのヒアリングを通して、実ユーザーの行動や選択をシミュレーションする技術および基盤（以下、本技術）を開発しました。本技術を活用し、さまざまなマーケティング施策を事前にシミュレーションすることで、小売店舗などを運営する企業は新商品企画や需要予測などを効果的に行うことが可能となります。

LLM は、膨大なデータをもとに学習しており、世界中のさまざまな人口統計的な知識や性格に関する情報を内包しています。そのため、一般的な属性（性別、年齢層など）や職業などのプロフィール情報を与えて質問すると、そのプロフィールに基づいた回答を簡単に得ることが可能です。この特性を利用するため、LLM はゲームキャラクターの対話生成やマーケティングのデプスインタビューなどに活用され始めています。一方で、汎用的なプロフィールや単純に過去のデータを与えた場合に得られる回答は、一般的な回答や過去データに含まれる単語に強く影響され、示唆に乏しく、また、その回答内容の信頼性を定量評価しにくい課題がありました。加えて、LLM は複数の選択肢について前方を好む傾向（ポジションバイアス）や、学習データセットに依存する知識のバイアスにより、期待通りの回答を得ることが難しいという課題もありました。

本技術は、バイアスの影響を抑えるように学習したパラメーター（属性ベクトル）でユーザーモデルを生成し、さらに数千～数万の質問を大量に並列実行した結果を統計的にまとめることで、LLM が内包するバイアスを取り除き、精度と信頼性の高い回答を取得します。

ユーザーモデル生成においては、匿名化された企業固有のデータをさらに多段階のクラスタリングで統計化したクラスタ情報をもとに出力データセットを生成し、属性ベクトルの抽出に利用します。属性ベクトルを LLM に加えることで、さまざまな属性（性格、価値観、考え方など）の付与が可能となり、この属性をユーザーモデルのプロフィールとして利用します。

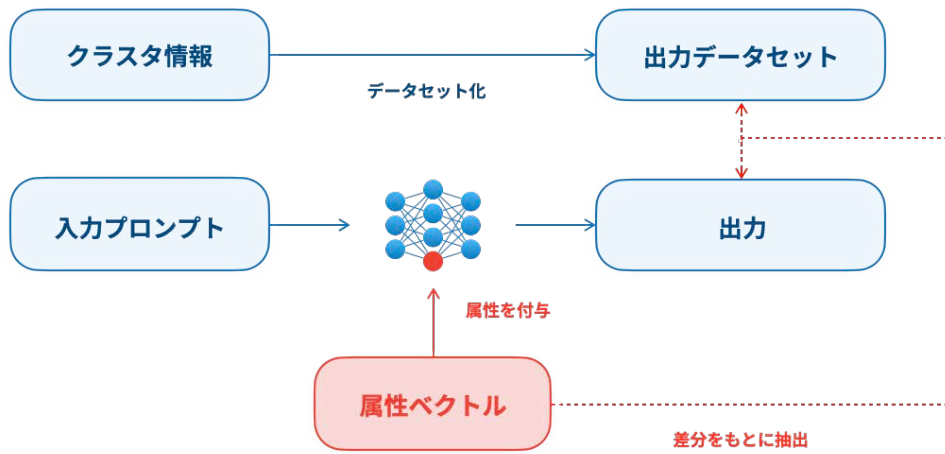


図 1. 属性ベクトルの抽出

また、本技術では大量の質問を短時間で処理するために、属性ベクトルの内部計算や効率的なキャッシュ処理を行う高速な LLM サービング基盤を開発し、従来手法より回答生成において最大 6 倍^{※1}の計算速度の向上を達成しています。

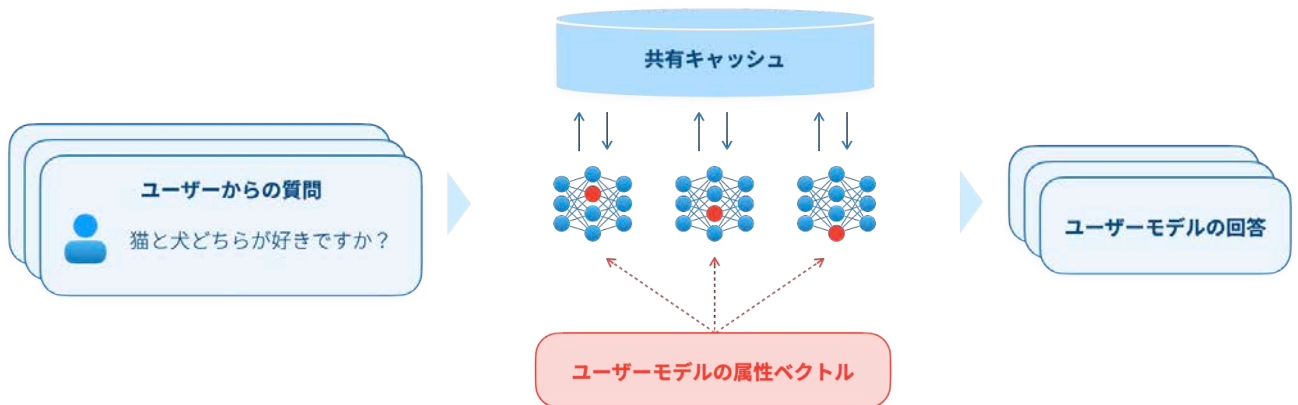


図 2. LLM サービング基盤の仕組み

本技術の効果を検証するため、ドコモが独自に実施した味覚に関する 2 万人のアンケートを用いて、各設問に対して 2 万人の平均に近い回答を生成できるかどうかの評価実験を行いました。



図 3. 実験内容

実験の結果、通常の LLM を用いた場合、最小 30%、最大 65%と正答率が安定しなかった結果に対し、属性ベクトルを与えた上で質問をパラフレーズ（質問の順序入れ替えなど）した本技術において 73%の安定した正答率の結果となり、技術の有効性を確認しました。

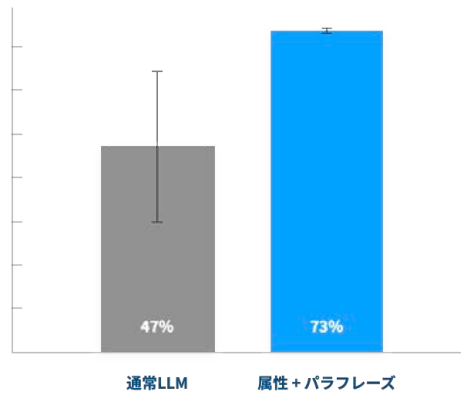


図 4. 正答率の比較

さらに、小売店舗で利用者の好みに合った新商品の選定に関する実証実験として、株式会社 Relic が運営する小売店舗^{※2}において、発注担当による選定、一般的な LLM による選定、本技術を活用した選定の 3 つの手法を用いて、商品の販売率を比較したところ、本技術は発注担当による発注と同等の販売率となり、一般的な LLM による選定よりも販売率が約 4% 上回る結果を確認しました。

また、NTT コミュニケーションズ株式会社が企画・運営を行うコンビニエンスストア^{※3}での実証実験において、本技術は新商品の売上順位を高精度に予測することを確認しました。

本技術は、目的に応じた適切なユーザーモデルを生成することにより、市場調査や AB テスト、商品開発などのさまざまなユースケースでの活用が期待できます。

売上順位	実績	予測
1	コーヒー ブラック	コーヒー ブラック
2	緑茶 A	炭酸飲料
3	炭酸飲料	緑茶 B
～	～	～
15	ハーブティー	新商品飲料
16	新商品飲料	炭酸水
17	フルーツジュース A	ハーブティー
18	緑茶 C	ほうじ茶
19	フルーツジュース B	天然水
20	天然水	緑茶 C

表 1. 新商品の売上順位予測

ドコモは本技術の自社サービスへの適用を検討するとともに今後さまざまな企業と連携し、さらなる精度検証を進め、技術の高度化、汎用化および社会実装を行い、お客さまのニーズや潜在的な要望をより深く理解し、それに基づいた最適な商品・サービスを提供することで、お客さま体験の向上をめざしてまいります。

※1 LLM プロバイダーを単体コールした場合と比べて 1,000~2,000 倍、高速推論エンジンの OSS と比較し最大 6 倍高速です。

※2 本技術の検証店舗に、株式会社 Relic が運営する「コンビニエンススタンド TUKTUK」の一部店舗を利用しました。

※3 本技術の検証店舗に、NTT コミュニケーションズ株式会社が本社大手町プレイス 28 階に設置している無人物販店舗を利用しました。

本件に関する報道機関からのお問い合わせ先

株式会社 NTT ドコモ

R&D 戦略部 社会実装推進担当

TEL : openinov-rd-ml@ml.nttdocomo.com

開発技術の概要

本技術は、図 5 の概要図の通り大きく 2 つのステップ、①ユーザーモデル生成、②確率的な予測、で構成されます。

① ユーザーモデルの生成

② 確率的な予測

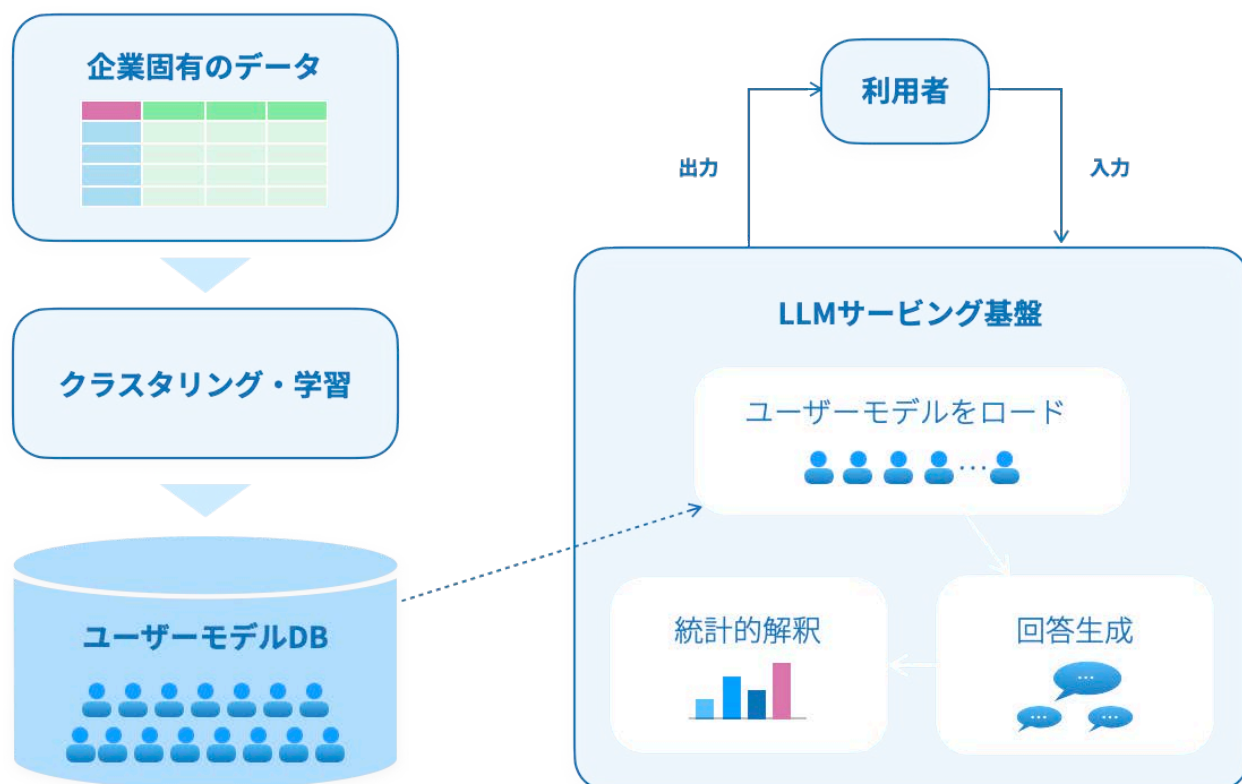


図 5. 本技術の概要図

①ユーザーモデル生成は、企業が持つ匿名化された固有のデータからユーザーモデルを生成します。企業固有のデータを多段階のクラスタリング技術で処理し、得られたクラスタ情報を用いてユーザーモデルの属性ベクトルを学習します。属性ベクトルは、学習の過程で趣味嗜好や価値観、デモグラフィック情報などの属性を獲得する数値列で、主に LLM の内部アクティベーションに直接足し合わせることでユーザーモデルのプロファイルを LLM に与えます。また、属性ベクトルは加減算可能なため、複数の属性ベクトルを組み合わせることで計算し、バイアスに強い、より純化された属性ベクトルに変換することも可能です。

②確率的な予測においては、生成したユーザーモデルに対して、予測対象に関するヒアリングを実施し、その回答を実ユーザーの行動予測とします。ユーザーモデルから信頼性の高い回答を得るために、属性ベクトル技術やキャッシュ処理を組み込み、並列性能を高めた LLM サービング基盤上で処理を行います。この基盤により、ユーザーモデルに大量の質問を行うことや、任意の集団に属する大量のユーザーモデル群に対し質問することが可能となり、回答結果を統計的に解釈することで、予測精度の向上を実現しています。