トピックス

2026 年 9 月 28 日
株式会社 NTT ドコモ

分析・予測対象の過去データが不足する環境でも 高精度な予測を実現する AI 技術を確立

～推薦システム分野の国際会議「ACM RecSys 2026」に採録～

株式会社 NTT ドコモ（以下、ドコモ）は、分析・予測対象の過去データが不足する環境でも高精度な予測を実現する AI 技術として Dual-view Adaptive Retrieval augmented Tweedie model（DART［ダート］手法、以下、本技術）を確立しました。本技術は、新規サービスの開始時など、予測に必要な過去データが不足し、AI が適切な判断を行えない「コールドスタート問題※1」を解決するもので、今後屋外デジタル広告（DOOH）※2 をはじめとするさまざまな分野の新規事業での活用が期待されます。

なお、本技術に関する論文が、推薦システム分野で世界的に権威のある国際会議「ACM RecSys 2026」※3 に採録され、その新規性と性能の高さが国際的に認められました。

新規サービスを立ち上げる際や、既存サービスを異なる領域や地域へ展開した直後においては、AI の予測に必要な過去データが十分に蓄積されていません。そのため、AI によるおすすめ（レコメンド）機能や広告の効果予測などが適切に機能しない「コールドスタート問題」が生じます。これにより、利用者の関心に合わせたおすすめや広告を提示できず、利用者の定着や事業の立ち上がりに時間を要する事例が発生していました。

本技術は、このコールドスタート問題を解決するために開発された新しい AI 技術です。主に 2 つの特長により従来の AI では対応が難しいとされていたデータが極端に少ない状況でも高精度な予測を可能にします。

1 つ目は、多様なデータ分布を柔軟に扱える「Tweedie（ツイーディ）分布※4」を AI の学習ルールに導入した点です。従来の AI の学習ルールは、平均値の周辺にデータが集まることを前提とした「ガウス分布※5」が広く採用されていましたが、SNS 上の反応のようにピーク時と非ピーク時で値が大きく異なるデータやばらつきがあるデータなど、複雑な条件下においては、AI が十分に学習できないという課題がありました。本技術では、そのような複雑な条件下でも柔軟に学習できる「Tweedie 分布」を採用し、データに偏りがある状況でもより適切な学習と予測を実現します。

2 つ目は、予測対象に類似した他の事例である「近傍事例」の活用です。予測対象の過去データが不足する場合でも、同じ地域の別の店舗や、類似するジャンルの別の商品などの近傍事例のデータから、場所・時間・属性といった共通する複数の特徴を自動で探して学習することで、高精度の予測を実現します。

本技術の活用により、サービス提供者は自社サービスにおける利用者のデータが集まるのを待つことなく、サービス開始初日から AI による高精度な予測やレコメンド機能を提供できるようになり、事業の早期立ち上げにつながります。例えば、DOOH 事業においては、都心の主要駅など人流が大きく変動する場所に新設したデジタルサイネージについて、広告がどのくらいの人に見られるか（インプレッション※6）を、過去データが少ない段階でも初日から高精度に予測できます。これにより、インプレッション数に応じた広告枠の価格を早期に設定し、販売開始までの期間を短縮できます。また、利用者はサービス利用開始時点から自身の興味関心に合ったコンテンツや情報を受け取りやすくなります。

ドコモは、2026 年度中に国内外の DOOH 事業で本技術の実証実験を実施する予定です。今後、実証を通じて有効性を検証し、グローバルでの本技術の活用をめざします。

※1 サービス開始初期は分析・予測対象の過去データが不足するため、AI が適切な予測や推薦を行えないという問題。「データゼロ問題」とも呼ばれる。

※2 Digital Out of Home の略。交通広告、屋外広告、商業施設などに設置されたデジタルサイネージ（電子看板）を活用した広告媒体。

※3 コンピュータサイエンス分野の国際学会 ACM が主催する、推薦システムに関するトップレベルの国際会議（The 20th ACM Conference on Recommender Systems）。

※4 統計学で用いられる確率分布の一種。ゼロ（0）の値が多く含まれるデータや、値のばらつきが大きいデータなど、現実世界に多く見られる複雑なデータ分布を柔軟に表現できる特長を持つ。

※5 平均値周辺が一番高くなり、左右対称に山のような形を描く分布。「平均に近いデータが多く、平均から離れるほどめったに起きない」という自然なデータの集まり方を表す。

※6 広告が利用者の画面に表示された回数のこと。

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 クロステック開発部　モビリティデータ活用技術開発担当 E-Mail: xt_md_marketing-dtc@ml.nttdocomo.com

別紙

「Dual-view Adaptive Retrieval augmented Tweedie model （DART 手法）」の概要

1. 概要

DART 手法（以下、本技術）は、AI による分析・予測に必要な過去データが不足する状況（コールドスタート）でも高精度な予測を実現する AI 技術です。

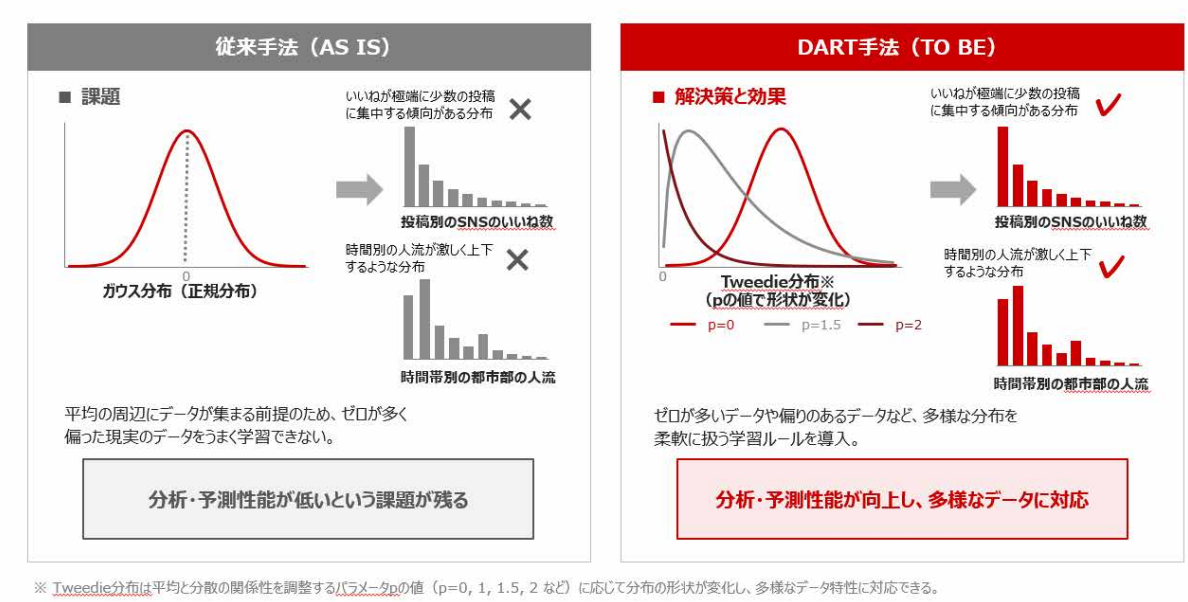
「Tweedie 分布」の導入による学習ルールの変更と「近傍事例」の活用という 2 つの特長を組み合わせることで、従来の AI では分析・予測が難しい環境に対応します。

2. 本技術の特長

（1）Tweedie 分布の導入による学習ルールの変更

従来の AI 手法の多くは、平均値の周辺にデータが集まることを前提とする「ガウス分布」を学習ルールに採用しています。しかし、現実世界のデータは、SNS 上の「いいね」数や都市部の人流のように、「ほとんどゼロだが、ごく稀に大ヒットが生まれる」「時間帯によって大きな変動がある」といった偏った分布を示すことが多くあり、ガウス分布ではそれらの偏った分布のデータをうまく学習できないという課題がありました。

このような課題を受け、本技術では、ゼロが多いデータや偏りのあるデータなど、現実世界の多様なデータ分布を柔軟に扱うことができる「Tweedie 分布」を AI の学習ルールに導入しました。これにより、偏りのあるデータでも AI での学習が可能となり、AI モデルの分析・予測性能の向上を実現しています。



※ Tweedie分布は平均と分散の関係性を調整するパラメータpの値（p=0, 1, 1.5, 2 など）に応じて分布の形状が変化し、多様なデータ特性に対応できる。

（2）近傍事例の活用

予測したい対象の過去データが不足する場合でも、同じ地域の別の店舗や似たジャンルの別の商品など、その対象に類似した近傍事例のデータから、場所・時間・属性といった複数の共通した特徴を自動で探し出して学習する「近傍検索拡張手法」を活用することで、精度の高い予測を実現します。