

大規模言語モデルにより事前学習不要なアート情報の自動分類が可能に

機械学習モデルを活用した情報の自動分類を行う際に、これまでタスクに合わせて都度データを再学習させ直す必要がありました。本研究では、アート領域のデータについて、大規模言語モデルを用いることで、新規データを学習させなくとも十分な精度で自動分類可能なことを明らかにしました。

近年、アート作品が投資の対象として注目を集め、その価格を予測することへの関心が高まっています。しかし、価格予測に必要なデータの整理（アノテーション）には、多くの人的資源と長時間の作業を要することが課題でした。そこで本研究では、大規模言語モデル（LLM）を使い、あらかじめ用意した学習データがなくてもデータを分類できる「ゼロショット分類」という手法を、アート作品の分類に適用することを試みました。

オープンモデルとして公開されている大規模言語モデル「Llama-3 70B」を4ビット形式に軽量化し、アート作品の種類（絵画、版画、彫刻、写真など）を自動判定できるかを調べました。その結果、90%以上の高い精度で、アートの種類を分類できることが確認されました。また、OpenAI社の生成AI「GPT-4o」と比べたところ、わずかながら高い精度を示しました。

本手法により、これまでの機械学習手法と同等の性能を、はるかに少ない人間の労力で実現できるため、アート情報に関するデータ整理の手間が大幅に削減できます。このような成果は、アート作品の多様な分析や価格評価を身近にし、投資だけでなく研究や鑑賞の面でも新たな活用の道が開けると期待されます。

研究代表者

筑波大学ビジネスサイエンス系

吉田 光男 准教授

戸嶋 龍哉 リスク・レジリエンス工学学位プログラム 博士後期課程1年

研究の背景

近年、アート作品が投資の対象として注目され、価格の予測が重要視されています。特に、富裕層の間では、「情熱投資」としてアートに資産の一部を投じることが一般的な投資活動となっており、アートの金銭的な価値や予測リターンを計算するために、価格予測モデルを構築するニーズが高まっています。これまで、アートの価格予測には統計モデルや機械学習^{注1)}が使われてきましたが、モデルを構築するためには大量のデータを整理（アノテーション）する必要がありました。この作業には多くの人手と時間がかかるため、より効率的な方法が求められています。

研究内容と成果

本研究では、アート作品の価格予測に必要なデータの整理を効率化するために、大規模言語モデル（Large Language Model: LLM）^{注2)}を活用した「ゼロショット分類^{注3)}」という手法の適用を提案しました。この手法は、新規の学習データを使わずに、未知のデータに対して自動的に分類を行うもので、膨大な手作業を必要とせず、アート作品の特徴を迅速かつ正確に分類することができます。

実験では LLM として、オープンモデルとして公開されている Meta 社の Llama-2 と Llama-3、Google 社の Gemma、そして、クローズドモデルである OpenAI 社の GPT-4o、GPT-4o-mini を採用しました。オープンモデルに対しては、データを圧縮して扱いやすくするために「4 ビット量子化」という技術を適用し、一般的な計算機でも扱いやすいように軽量化しました。これにより、ローカルサーバでの実行が可能になります。これらの LLM を使用した手法と、これまでよく用いられていた機械学習を利用した手法とを比較しました（参考図）。

具体的なタスクとして、アート作品の種類（絵画、版画、彫刻、写真など）を自動で分類する実験を行いました。この際、作品の材質や技法を記載した「medium（メディウム）」と呼ばれる情報や作家、作品サイズ情報、オークション売買情報などをもとに、作品の種類を分類しました。例えば、「Banksy（バンクシー）」という作家の「Girl with Balloon, 2004」という作品名で、「Screenprint in colours, on wove paper（網目紙にカラーで孔版印刷された版画）」というメディウムを持ち、サイズは「65.71 x 49.83 cm」で、オークションハウス「Bonhams（ボナムズ）」の「Prints and Multiples」というオークションに出品されています。この作品が「Print（版画）」に分類されれば正解となります。

実験の結果、Llama-3 70B モデルは、90%以上という高い精度でアート作品の分類を行うことができました。この精度は、商業的に利用されている OpenAI の GPT-4o モデルと比較してもわずかに高く、従来の機械学習手法とも同程度の性能です。Llama-3 はオープンモデルで無料で利用できるため、独自のシステムなどにも組み込みやすいという利点があります。本研究成果は、LLM を用いたゼロショット分類により、アート市場におけるデータ整理の負担を大幅に軽減し、価格予測モデルの構築のためのアノテーションに関する人的コストを下げられる可能性を示しました。

今後の展開

本手法は、アート情報の分類以外の整理作業に適用できる可能性があります。例えば、収集したデータのフォーマット変換、作品のジャンルの推定などです。また、このフレームワークは他の分野にも応用可能であり、さまざまな複雑なデータに対する統一的な自動前処理手法の開発も期待されます。今後、さらなる精度向上のために、指示の最適化（プロンプトエンジニアリング）が有効だと考えられます。これには、プロンプト自動生成手法や少数カテゴリに対する分類の偏りを抑える手法の開発が求められますが、評価用データの作成は高コストという課題があるため、総合的なアプローチが必要です。

参考図

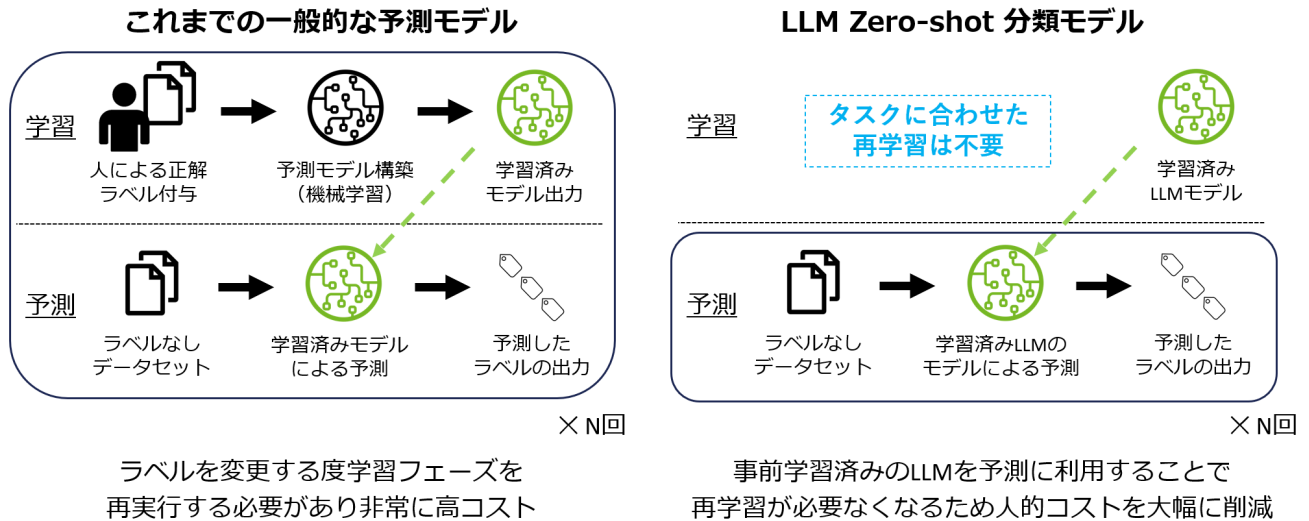


図 本研究の概要

従来の予測モデル（左図）では、対象ごとに学習データとラベルのペアを人手で作成し、対象ラベルを自動で付与する機械学習を利用した予測モデルを構築し、新規データの予測を行う。この手法はラベルや対象を変更するたびに学習フェーズを再実行するため、予測モデルの変更にはコストがかかる。本研究（右図）では、事前学習済みの LLM を用いたゼロショット分類により、学習の手間を大幅に削減しつつ、アート情報の分類タスクにおいて、従来モデルとほぼ同等の分類精度を達成可能なことを実証した。

用語解説

注1) 機械学習 (Machine Learning: ML)

コンピュータがデータを学習し、自動でルール、パターンや関係性を見つける技術。

注2) 大規模言語モデル (Large Language Model: LLM)

非常に大きなパラメータ数を持つ言語モデル。構築のためには大量のデータを学習させる必要がある。自然でまとまりのある文章を生成でき、文章の読解、翻訳、質問応答、要約などの分野で、特定のタスク向けに追加の学習をしなくても、幅広い用途に高性能で対応できる。

注3) ゼロショット分類 (Zero-shot Classification)

機械学習において、すでに獲得した知識のみを用いて、追加の学習を行うことなく新規ラベルの分類を行う分類タスクのこと。

研究資金

(該当なし)

掲載論文

【題名】 Zero-Shot Classification of Art With Large Language Models
(大規模言語モデルを用いたアートに関するゼロショット分類)

【著者名】 Tatsuya Tojima, Mitsuo Yoshida

【掲載誌】 *IEEE Access*

【掲載日】 2025年1月23日

【DOI】 10.1109/ACCESS.2025.3532995

問合わせ先

【研究に関すること】

吉田 光男（よしだ みつお）

筑波大学 ビジネスサイエンス系 准教授

URL: <https://www.gssm.otsuka.tsukuba.ac.jp/~mitsuo/>

【取材・報道に関すること】

筑波大学広報局

TEL: 029-853-2040

E-mail: kohositu@un.tsukuba.ac.jp