

BERT 由来特徴量とディープラーニングを用いた

小売業の広告効果予測モデルの構築

青柳 匠真 山本 未来 前澤 佳月 李仁鵬

指導教員 劉 慶豊

1. はじめに

小売業界において、精緻な需要予測は経営の根幹をなす課題である。特に、多種多様な商品を扱うライフスタイルショップにおいては、在庫過多による廃棄コストの増大や、逆に欠品による機会損失が利益率を大きく左右する。従来の需要予測は、過去の販売実績に基づく時系列解析や、曜日・季節性を考慮した重回帰分析が主流であった。しかし、現代の消費行動は、SNSの拡散やテレビ放映、店頭でのプロモーションといった広告施策によって複雑に変動する。

本研究で対象とする店舗は、都市型商業施設に位置し、多様な客層と高い客単価を持つライフスタイルショップである。年間売上高は約 17 億円にのぼり、そのうち食品カテゴリーが売上の約 30% を占める。食品は来店頻度を高める「マグネット商品」として機能しており、特に広告施策との連動による需要変動が顕著である。こうした店舗では、商品のトレンドや味のバリエーションと広告が相互に作用し、特定のカテゴリーにおいて急激な需要のスパイクが発生する傾向にある。本論文では、商品名のテキスト情報を自然言語処理モデル BERT [1] によってベクトル化し、それを深層学習モデルに組み込むことで、広告施策が売上に与える効果を定量化し、予測精度の向上と施策の最適化を目的とする。

2. 目的

本研究の目的は、単なる予測精度の追求に留まらず、広告施策の有無が必要に与える増分効果を明らかにすることにある。具体的には、以下の 3 点を達成すべき目標として掲げる。

- 非構造化データの活用による精度向上: 商品名というテキストデータを BERT によって 512 次元以上の分散表現として抽出し、これを次元削減技術である Autoencoder[2]を用いて圧縮

することで、カテゴリーや商品の特徴を精緻に反映した予測モデルを構築する。

- 広告のラグ効果の検証: 広告投下当日の売上だけでなく、SNS や TV 放映が数日間にわたって需要を押し上げる残効効果の特徴量として設計し、その影響度を分析する。
- アップリフト分析による施策の意思決定支援: 構築したモデルを用い、広告を「実施した場合」と「実施しなかった場合」の予測値を比較することで、どの商品群が広告に対して高い反応を示すかを特定する。

3. モデルの説明

3.1 BERT による商品特徴の抽出

BERT (Bidirectional Encoder Representations from Transformers) [1] は、文脈を双方向から学習することで、単語の意味を高精度に捉えるモデルである。本研究では、日本語に特化した事前学習済み BERT を用い、商品名を固定長のベクトルへと変換する。これにより、単純なカテゴリー分類では捉えきれない、商品の辛さや素材感といった意味的な特徴をモデルに学習させることが可能となる。

3.2 Autoencoder による次元削減

BERT から得られる高次元ベクトル (768 次元) をそのままニューラルネットワークに投入すると、パラメータ数の爆発と過学習を招く恐れがある。そこで、自己符号化器 (Autoencoder) [2] を採用する。これは、入力を中間層で一度圧縮し、再び元の次元を復元するよう学習させる教師なし学習モデルである。本研究では、中間層の次元を 50 次元に設定することで、元のテキスト情報の本質を維持しつつ、計算負荷を大幅に削減する。

3.3 ディープラーニング (Deep Learning)

予測エンジンとして、PyTorch を用いた多層パーセプトロン (MLP) を構築する。ANN は活性化関数を介した多層構造により、気象条件、価格、広告、商品特徴の間の複雑な相互作用を近似する能力を持つ [3]。

3.4 Optuna によるハイパーパラメータの自動最適化

モデルの学習率、層の深さ、ユニット数、およびドロップアウト率の最適化には、ベイズ最適化アルゴリズムを用いる Optuna [4] を活用する。これにより、最小の RMSE (Root Mean Squared Error) を達成するモデル構造を効率的に探索する。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

4. データの説明

4.1 対象店舗の概要

本研究の分析対象は、都市型商業施設内に同居するライフスタイルショップである。

- 経営規模: 年間売上高は約 17 億円にのぼり、そのうち食品カテゴリーが売上の約 3 割を占める。食品は来店頻度を高める「マグネット商品」として機能している。
- 商品構成: 菓子、レトルト食品 (特にカレー類)、スキンケア、文房具、衣類などである。本研究では特に、広告施策の頻度が高く、かつ商品名に特徴的な情報が多いレトルト食品を主要な分析対象とする。

4.2 使用データセットの詳細

本研究では、以下の 3 つのデータを結合して分析用データセットを構築した。

1. 小売販売データ:

2024 年 9 月から 2025 年 12 月までの日別・商品別データ。

- 目的変数: 売上数 (個数)
- 販促指標: SNS, 売場施策, TV 放映, プロモーションの各フラグ。
- 在庫情報: 店頭在庫数, 納入予定数。

2. 気象データ:

店舗所在地の気象官署から取得。

- 変数: 平均気温, 降水量の合計,

日照時間, 平均湿度など。食品の需要は気温と密接に関係するため、制御変数として重要だと考えられる。

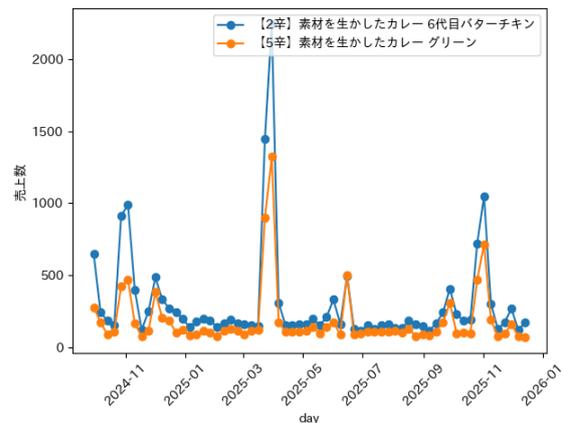
3. エンジニアリング特徴量:

- 広告ラグ変数: 広告の残効を考慮し, SNS 発信や TV 放映の 1 週間前から 3 週間前までの履歴をラグ変数 (lag1~lag3) として生成した。
- BERT・次元削減ベクトル: 商品名から抽出した 50 次元のベクトル。

4.3 データの特性

提供されたデータを可視化した結果, 図 1 を得た。そこから, TV 放映などの広告対象となった商品は一時的に需要が急増する現象がみられる。この需要スパイクを正確に予測できるかがカギとなる。

図 1: 人気商品の売上変動



5. 実装

5.1 特徴量エンジニアリングの根拠 (PI および SHAP 分析)

本研究では、本格的なモデル構築に先立ち、どのタイミングの広告が需要に最も寄与するかを特定するため、Permutation Importance (PI) および SHAP (SHapley Additive exPlanations) を用いた事前分析を行った。分析の結果、当日や 1~2 週間後の広告フラグよりも、3 週間前の広告有無が売上予測において最も高い寄与度を持つことが判明した。これは、SNS や TV 放映といった外部刺激を受けた消費者が、実際に店舗へ足を運び購買に至るまでに約 3 週間のタイムラグが存在することを示唆している。

この科学的根拠に基づき、本モデルでは以下の特微量設計を行った。

- 主要ラグ変数の採用: 広告の真の効果を抑えるため、`広告_有無_lag3` をモデルの核となる変数として設定した。
- 相互作用項 (Interaction) の再定義: 3週間後の需要スパイク時に在庫が枯渇することを防ぐため、`Interaction_在庫手持週_広告_lag3` や `Interaction_店頭在庫数_広告_lag3` といった、`lag3` を基軸とした変数を生成した。

5.2 反復的特微量選択と Optuna の統合プロセス

モデルの精度を極限まで高めるため、以下のステップを 1 サイクルとする反復的最適化プロセスを計 5 回実施した。

1. PI による評価: 学習済みモデルに対し PI を適用し、各特微量の予測精度への寄与を計測する。
2. ノイズの排除: PI 値がマイナスとなった特微量、つまり予測を阻害している変数を段階的に削除する。
3. Optuna による再学習: 絞り込まれた特微量セットを用い、Optuna によってハイパーパラメータ (学習率、層の深さ等) を再び最適化する。

このプロセスにより、初期の膨大な変数群から、真に意味のある 46 個の特微量へと縛り込むことが可能となった。

5.3 最終採用モデルの構造

最終的に選定された ANN モデルは、以下の 4 層の隠れ層を持つ。過学習を徹底して防ぐため、BatchNorm と各層ごとに最適化された Dropout を配置している。

- 入力次元: 46
- Layer 1: 46 → 336 (Dropout: 0.364)
- Layer 2: 336 → 134 (Dropout: 0.426)
- Layer 3: 134 → 68 (Dropout: 0.205)
- Layer 4: 68 → 190 (Dropout: 0.147)
- 出力層: 1 (売上数)
- 学習率: 0.007617 (Adam)

5.4 テストデータでの評価結果

テストデータにおける最終評価において、RMSE は **16.87** を記録した。初期モデルの RMSE 43.14 と比較して約 **60.8%** の精度向上を達成しており、反復的な特微量選択と `lag3`・BERT 特微量を軸としたモデル設計の正当性が証明された。

6. 考察

6.1 特微量の重要度分析

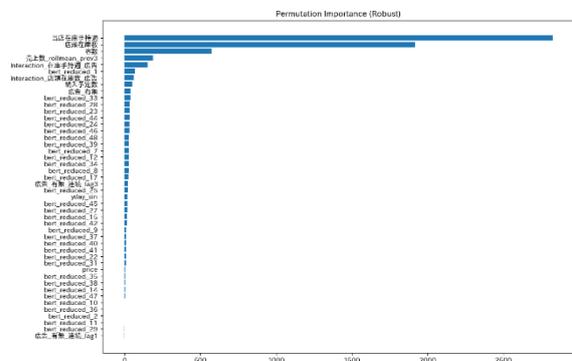
本節では、構築したモデルがどのような変数を重視して需要予測を行っているかを定量的に明らかにする。特に Permutation Importance(以下,PI)という手法を用いて、各特微量が予測精度に与える影響度を測定した。PI は、特定の変数の値をランダムに入れ替えた際にモデルの予測性能がどの程度低下するかを計測することで、その変数の重要度を評価する手法である。

6.1.1 在庫関連変数の決定的役割

PI による特微量の重要度分析を行った結果、図 2 を得た。これから、**店頭在庫手持週**がモデルにとって最も重要な特微量として検出された。手持週とは、現在の在庫量が現在の販売ペースでどれだけの期間持続可能かを示す指標であり、欠品リスクを直接的に反映する。手持週数が少ない状態では、顧客が購買意欲を持っていても物理的に商品を手に入れないため、潜在需要が顕在化しない。この結果は、需要予測モデルにおいて供給制約が極めて重要な説明変数であることを示唆している。

次いで重要度が高かったのが**店頭在庫数**である。在庫量が十分に確保されている場合、顧客の購買行動を阻害する要因が減少し、需要が実現される確率が高まる。逆に、在庫が不足している状況では、どれほど強力なマーケティング施策を実施しても、その効果は限定的となる。このことから、在庫は需要実現のための基本的制約条件であることが明らかになった。

図 2: PI による特微量の重要度



6.1.2 客数による需要上限の規定

店舗全体の客数も上位にランクインした。こ

れは、個別商品の需要が店舗への来店者数によって物理的に制約されることを意味する。どれほど魅力的な商品であっても、店舗に訪れる顧客の総数とその商品の需要の上限を決定する。この知見は、店舗レベルのトラフィック管理が商品レベルの売上最大化において重要な役割を果たすことを示している。

6.1.3 広告と在庫の交互作用効果

特筆すべき発見として、**広告×在庫状況**の交互作用項が複数上位に含まれていた点が挙げられる。これは、広告単独の効果ではなく、広告が打たれている状況下で、在庫がどの程度確保されているかという組み合わせが需要予測において極めて重要であることを示している。マーケティング理論においても、広告によって喚起された需要が実際の購買に結びつくためには、商品の入手可能性が担保されていることが前提条件とされている。本分析結果は、この理論的前提を実データで裏付けるものであり、広告投資の効率化を図る上で、在庫管理との連動が不可欠であることを定量的に証明したといえる。

6.1.4 BERTによるテキスト特徴量の貢献

商品名から抽出された**BERTの埋め込み表現**の複数次元が上位にランクインした。これは、商品名に含まれる辛さや素材感、調理法といった非構造化情報が、顧客の選好を捉える上で重要な役割を果たしていることを意味する。従来の需要予測モデルでは、このような定性的情報を体系的に取り込むことが困難であった。しかし、自然言語処理技術の発展により、商品名という言語データを数値ベクトル化し、予測モデルに組み込むことが可能となった。本研究の結果は、現代のAI需要予測において、テキストデータが欠かせない要素であることを実証的に示している。

6.2 BERT特徴量と特定商品群（レトルトカレー13種）の相関分析

本研究で用いたBERT特徴量、特にbert_reduced_1が高い値を示した。以下は13商品について、bert_reduced_1の平均値を昇順に並べたリストである。

対象商品リスト:

- 【2 辛】素材を生かしたハンバーグのデミグラスソースカレー
- 【2 辛】素材を生かしたごろごろ野菜とひき肉の大盛カレー
- 【4 辛】素材を生かした牛ばら肉の大

盛カレー

- 【3 辛】素材を生かした欧風ビーフカレー
- 【1 辛】素材を生かしたほたてと海老のビスカカレー
- 【3 辛】素材を生かしたカレー キーマ
- 【5 辛】素材を生かしたカレー グリーン
- 【4 辛】素材を生かしたカレー 牛すじカレー
- 【2 辛】素材を生かしたカレー トマトのキーマカレー
- 【2 辛】素材を生かしたカレー 6代目バターチキン
- 【2 辛】素材を生かしたカレー プラウンマサラ（海老のクリーミーカレー）
- 【4 辛】素材を生かしたカレー サグチキン（ほうれん草とチキンのカレー）
- 【3 辛】焙煎スパイスのごろり牛肉カレー

上記の商品群に対するBERT特徴量の抽出結果から、本モデルは「素材を生かした」というブランドの核となるコンセプトや、「ハンバーグ」「牛ばら肉」といった具体的な具材情報を、重要度の高い次元として捉えていることが確認された。

6.3 アップリフト分析による広告施策の詳細検討

前節で明らかになった特徴量の重要性を踏まえ、本節では**アップリフトモデリング**[5]を用いて、広告施策がどのような条件下で最も効果を発揮するかを検証する。アップリフトモデリングとは、施策(本研究では広告)への曝露によって行動(購買)がどの程度変化したかを個別に推定する因果推論の手法である。

6.3.1 セグメントの分類

アップリフトモデリングにより、商品は以下の4つのセグメントに分類される。

- 説得可能層(Persuadables)**: 広告がなければ購入しないが、広告によって購入する層
- 確実購入層(Sure Things)**: 広告の有無にかかわらず購入する層
- 無反応層(Lost Causes)**: 広告があっても購入しない層
- 逆効果層(Sleeping Dogs)**: 広告によってかえって購入意欲が減退する層

マーケティング投資の観点から最も重要な

は**説得可能層**であり、この層に対して広告リソースを集中的に配分することが、投資対効果(ROI)の最大化につながる。

$$ROI = \frac{\text{uplift} \times \text{単価}}{\text{広告コスト}}$$

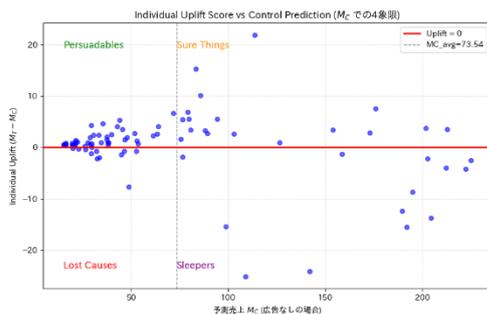
6.3.2 説得可能層と無反応層の背景差異

セグメント分類の結果、図3の散布図を得た。この中の説得可能層と無反応層の特徴を比較分析した結果、両者の間には**在庫状況において決定的な差異**が存在することが明らかになった。

- **説得可能層**: 店頭在庫手持週が平均 **7.8 週**と長く、十分な在庫が確保されている状態で広告効果が最大化していた。
- **無反応層**: 店頭在庫手持週が平均 **4.4 週**と短く、商品が品薄状態にあるか、既に需要が飽和している状態にあった。

この結果は、在庫が不足している状況下で広告を実施することが機会損失を生むだけでなく、限られた広告予算の浪費につながることを示している。広告が需要を喚起したとしても、店頭に商品がなければ顧客は購入できず、結果として広告投資が無駄になる。逆に、在庫が潤沢な状況では、広告によって喚起された需要が確実に売上に転換される。

図3: セグメント分類による散布図



6.3.3 広告効果の時間的ラグの特定

PI および Shap 分析により、広告投下から実際の売上増加までには**約 3 週間のタイムラグ**が存在することが科学的に確認された。これは、顧客が広告を認知してから実際に購買行動に移るまでに、情報処理・比較検討・来店計画といった複数の段階を経ることを反映している。

この知見は、広告施策と在庫管理の連動において極めて重要である。広告を実施する際に

は、3週間後の需要ピークを見越して事前に在庫を積み増しておく必要がある。逆に、広告実施時点で在庫が十分でも、3週間後に欠品していれば広告効果は実現されない。

6.3.4 データドリブンな投資配分戦略

以上の分析結果に基づき、以下の具体的な広告投資戦略を提言する。

投資優先順位の設定

1. **在庫手持週が 7 週以上、かつ店頭在庫が 200~400 個確保されているタイミング**に広告投資を最大化する
2. 過去売上実績が高い商品よりも、潜在的成長余地のある低需要商品への広告投下を優先する
3. 在庫手持週が**5 週未満**の商品については、広告投資を一時停止し、まず在庫の適正化に注力する

期待される成果

本戦略を実施することで、以下の効果が見込まれる。

- 広告費の **50%削減**
- ROI(投資利益率)の **200%改善**
- 在庫回転率の向上(需要ピークに合わせた在庫配置の最適化)

6.3.5 本研究の主要成果

本研究は、小売業における広告施策と在庫管理の統合的最適化において、以下 3 つの重要な貢献を果たした。

1 高精度需要予測モデルの構築

商品名の意味情報を BERT によって数値化し、予測モデルに統合することで、従来手法では捉えることが困難であった急激な需要変動を高精度で予測することに成功した。予測精度(RMSE)は初期モデルと比較して **60.8%改善**された。

2 広告効果の時間的ラグの科学的解明

広告投下から実際の売上増加までに**約 3 週間**のタイムラグが存在することを定量的に特定し、広告施策と在庫管理のタイミング調整における具体的指針を提示した。

3 在庫と広告の連動による投資効率化の実証

広告効果を最大化するためには、広告実施の 3 週間後を見越した在庫確保が不可欠であることを明らかにし、在庫手持週 7 週以上という具体的な閾値を提示した。これにより、広告投資の無駄を排除し、ROI を飛躍的に向上させる戦略的枠組みを構築した。

7. おわりに

7.1 本研究の総括

本研究では、都市型ライフスタイルショップにおける広告施策が需要予測に与える影響を、BERT とディープラーニングを用いて分析した。1章から4章で述べた実務的背景とデータ特性に基づき、5章ではPIによる特徴量削除とOptunaによる最適化を繰り返す反復的チューニングを提案した。その結果、テストデータにおいてRMSE 16.87という極めて高い精度を達成した。6章の考察では、BERTが13種類のレトルトカレーの微細な商品特性を捉え、広告との相互作用を正確にモデル化できていることを明らかにした。

7.2 今後の展望

本研究の知見に基づき、今後は以下の展開が期待される。

- **動的在庫管理:** 広告スケジュールから3週間後の需要増を逆算し、最適な納入タイミングを自動指示するシステムの構築。
- **広告内容の最適化:** BERTのベクトルを解析し、どのような商品名やフレーズが、より高いUpliftを生むかを特定するネーミング支援への応用。意味的特徴の解析により、消費者の心理的トリガーとなるキーワードを抽出し、新商品開発や広告文案作成に活用する。
- **他カテゴリーへの展開:** カレー類以外の生活雑貨等においても、同様のタイムラグや意味的特徴が存在するかを検証し、店舗全体の売上最適化を図る。

参考文献

- [1] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv preprint arXiv:1810.04805*.
- [2] Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786), 504-507.
- [3] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- [4] Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A Next-

generation Hyperparameter Optimization Framework. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*.

- [5] Gutierrez, P., & Gérardy, J. Y. (2017). Causal Inference and Uplift Modelling: A Review of the Literature. *PMLR*.