

2段階の機械学習予測モデルに基づく 季節性のある中古アパレル商品の需要予測に関する一考察

情報数理応用研究

5220C016-1 齊藤芙佑

指導教員 後藤正幸

A Study on Demand Prediction for Seasonal Second-hand Fashion Items Based on a Two-Stage Machine Learning Prediction Model

SAITO Fuyu

1. 研究背景と目的

近年のSDGsへの意識増大やコロナ感染症による在宅時間の増加に伴い、家中の不要品をECサイトやフリマアプリを用いて再販売する人が増加し、リユース市場は拡大傾向にある。特に、軽量で運搬がしやすく「古いからこそ個性と価値」を持つ中古アパレル商品市場の活性化は顕著である。一方、インターネットを介した売買により、購買・検索などの行動履歴情報など、中古品に関する膨大なデータの蓄積が可能となっている。これらを様々な機械学習手法と合わせて活用し、有効な施策に結びつける試みも数多くなされており、顧客分析や商品推薦などで成果を挙げつつある。

本研究で対象とするECサイトZOZOUSUEDは、顧客から中古アパレル商品の買取・販売のビジネスを展開している。ZOZOUSUEDでは、買取後すぐに出品される「通年商品」と、季節により需要の変動がある「季節商品」の双方が扱われており、季節商品である「春夏商品」と「秋冬商品」は買取後まず在庫となりシーズン単位で出品される。季節商品をシーズン序盤に過剰に出品してしまうとシーズン後半に出品可能な在庫がなくなり、ECサイト魅力度と顧客満足度が低下してしまう。一方、出品が過少になりすぎると、シーズン後半に商品を倉庫に保管したまま販売時期を逃し大量の売れ残り在庫を発生させてしまう。このようなトレードオフが存在するため、季節商品は通年商品と区別して、需要に沿って戦略的に出品されることが求められている。

従来では、在庫となった季節商品は、シーズンごとに出品管理者の属人的な予想に基づき出品されてきた。しかし、現状の属人的な予想は客観性に欠け、実績と大幅に異なってしまう場合もあり、上述のトレードオフの問題解決に繋がっていない可能性が考えられる。以上より、豊富に蓄積されたデータを活用することで、より定量的な判断に基づく的確な季節商品の出品計画を可能とする需要予測モデルの構築を検討する必要がある。

ここで、実ビジネスにおいては、予測時に得られない情報があったり過去に事例のない施策が行われたりする場合も多いため、外部要因を完全に網羅した予測は困難であると言える。よって、そのような一部の要因が考慮されていない予測結果をビジネス行動の自動化に結びつけることは非常にリスクが高いと考えられる。そこで、完全にモデルに頼る自動化を目指すのではなく、予測モデルと出品管理者の協働による予測の質向上を試みる。これにより、出品管理者は現場で予測結果を定量的な指標として使い、データ主導で意思決定を

行うことができる。なお、この方針を選択することは現場の出品管理者自身がモデルを利用することを意味する。よって、負担の少ないスムーズな導入のために、利用者の要求に十分に応えた従来方法に沿った設計が求められる。従来方法は、シーズン前とシーズン中の2段階の需要予想・出品計画である。出品管理者により、まずシーズン開始前に長期的なシーズン全体の需要を大まかに予想し商品の調達と在庫計画を行う。次にシーズン中のオペレーションとして日々の気象条件などの外部要因や直近の販売状況を確認し短期的に需要変動を予想し出品計画を策定する。そこで本提案手法も、以上のようなオペレーションに沿う長期予測モデルと短期予測モデルから成る2段階の需要予測手法とすることを考える。

以上より本研究では、季節商品の実ビジネスにおける定量的な出品計画の支援となる、長期予測モデルと短期予測モデルを組み合わせた2段階の需要予測手法を提案する。そして、提案モデルを実データに適用することで、提案モデルの信頼性と有用性の評価を行う。また、実ビジネスでの実証実験を設計・運用することで、本モデルが出品計画の支援に実際に有効であるかどうかを実証的に検証する。さらに、効果検証の観点から傾向スコアを用いた評価についても結果を示す。

2. ZOZOUSUEDのビジネスモデル

ZOZOUSUEDのビジネスモデルは、中古アパレル商品の買取・販売のECサイトを介した活動である。これにより蓄積されたデータの様々な施策への活用も積極的である。例として、適切な出品価格や価格保持期間の設定など、ミクロな視点から個々の商品に対する意思決定への支援等が挙げられる。

ZOZOUSUEDでは、季節商品に対し「春夏商品」と「秋冬商品」と呼ばれる商品群を独自に設定している。各商品群に対する販売数の推移を図1に示す。なお、縦軸の販売数に関しては、非公開情報であるため具体的な数値は伏せて表示している。このように、季節商品は季節性に応じて特徴的な需要変動を示す。さらに、春夏/秋冬商品は、より詳細に商品の特性を考慮した「弱/中/強」に細分化した商品カテゴリが設定されている。具体的には、春夏商品に関して、スプリングコートやニットカーディガンなどのより春らしい商品を「弱」、サンダルやノースリーブシャツなどのより夏らしい商品を「強」、それ以外の商品を「中」としている。同様に、秋冬商品に関して、ジャケットなどのより秋らしい商品を「弱」、ダウンコートなどのより冬らしい商品を「強」、それ以外の商品を「中」としている。これらは、その季節性の強

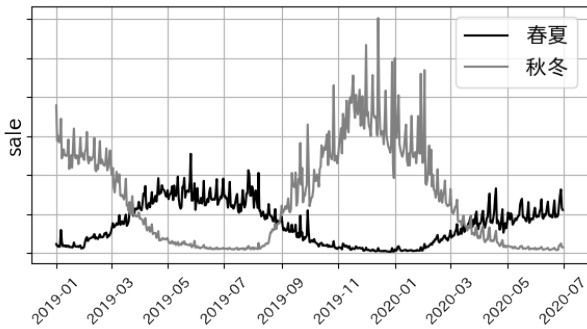


図 1: 2019 年度における販売数の分布

表 1: 商品カテゴリまとめ

	販売数割合	ピーク時期	特徴
春夏-弱	0.06	4月中旬/8月下旬	各ピークに向かって緩やかなカーブを描く
春夏-中	0.63	6月初旬	シーズン通して緩やかなカーブを描く
春夏-強	0.30	6月中旬	やや遅れたピーク、急なカーブを描く
秋冬-弱	0.08	10月初旬	ピークに向かって緩やかなカーブを描く
秋冬-中	0.64	11月初旬	シーズン通して緩やかなカーブを描く
秋冬-強	0.29	12月初旬	やや遅れたピーク、急なカーブを描く

さの違いにより、最も需要の高まるピーク時期やその前後の動向が異なる。各商品カテゴリの 2019 年の各シーズンにおける販売数とその割合、最も販売数が多いピーク時期、需要変動に関する特徴を表 1 にまとめる。

ここで、需要を的確に捉え日々適切な量を出品することは、出品過剰に起因した欠品による EC サイトの魅力度及び顧客満足度低下の問題と、出品過少に起因する販売時期見逃しによる売れ残りの問題防止のために非常に重要である。さらに、顧客が求める商品を適切なタイミングと量で出品することで、顧客満足度向上、売上増加、及び即時販売による在庫管理コスト削減が期待できる。そこで、ファッションビジネスにおいて重要な存在である季節商品の出品に対し、出品管理者が、上述の商品カテゴリごとで需要が高まるシーズンに以下の手順で需要の傾向に沿うように計画・管理している。

1. シーズン開始前：過去の経験やノウハウに基づいた予想によりシーズン全体の動向を見据え、シーズン全体のマクロな在庫計画と長期的な出品計画を策定する。
2. シーズン中：日々の販売状況の動向や外部要因の影響を確認しつつ、過去の経験やノウハウに基づき需要変動を予想し、短期的な出品計画を策定し実行する。

このように、従来の出品計画方法は、出品管理者の属人的な需要の予想に大きく依存すると言える。

3. 提案手法

3.1. 問題設定

本研究では、季節商品の出品計画の問題に焦点を当てる。従来の季節商品の出品は、出品管理者の経験やノウハウに基づいた属人的な需要予想により判断されてきた。しかし、この属人的な予想と実績に乖離がある場合、様々な問題が生じる。そこで本研究では、需要予測モデルを構築することで、

定量的な根拠に基づく客観的かつ確かな計画立案を支援し、モデルと人間の協働による意思決定の質向上を試みる。なお、ZOZOUSSED では、ミクロな視点で個別商品の意思決定を支援する数多くの機械学習モデルが検討されてきた。これに対し本研究では、これらの先行研究では考えられていなかった、マクロな視点で商品群に対する包括的な意思決定への支援に着目した研究を行う。

3.2. 提案への着想

本研究は実運用化を想定しているため、現場でモデルを利用する出品管理者の負担を軽減しスムーズな導入を促すことが望まれる。さらに実ビジネスへの適用条件には、リスク防止のため、十分な解釈性・高い精度の保持・迅速な結果の獲得が必須である。以上の理由から、本提案手法は商品カテゴリごとに構築された長期予測モデルと短期予測モデルから成る 2 段階構成とする。これにより、出品管理者は従来通り、商品カテゴリごとに 2 段階で計画を行うことができ、負担軽減およびスムーズな導入が期待できる。また、商品カテゴリごとにモデルを構築することで、季節性の強さの違いに由来する需要変動に関する特徴と要因との関係性をよりの確に捉えられ、予測の精度向上が期待できる。加えて、2 段階構成とすることで提案手法は加法モデルとなり、次のように精度と解釈性の向上を両立することも期待できる。まず、長期予測モデルにより短期予測モデルのバイアスを除去することで、精度向上が期待できる。さらに、短期予測モデルは長期予測モデルによる予測と実測値とのずれを新たに得られた複数の要因を考慮し修正する役割を担い、予測値に対する解釈性向上が期待できる。例えば、短期予測モデルがシーズン中の気象情報より負の予測値を出力し下方修正が行われた場合、「シーズン前の予測よりも、雨が多く気温が上がらなかったため、夏物の販売数量が伸びなかった」などの直感に沿った解釈が可能となる。

さらに、ZOZOUSSED に蓄積された過去の販売履歴データは、多様かつ膨大であることから、機械学習によるアプローチと相性が良いと考えられる。また、一般に顧客の中古アパレル商品の購入意向に影響を与える要因として、気温や天候などの環境条件や、価格割引率やセールイベントなどの EC サイトの条件が挙げられる。よって、実ビジネスへの適用の必須条件を満たしつつ、これら複数の外部要因を十分に考慮することも求められる。以上の理由から、本提案手法を構成するモデルの構築は、Light Gradient Boosting Machine (以下、LightGBM) アルゴリズム [1] を適用することにより行うことを考える。これは、多種多量の特徴量を考慮可能な、高精度かつ低計算コストなアンサンブル型回帰モデルで、本研究の理想とする需要予測モデルの構築が期待できる。

3.3. 提案手法

提案手法は、シーズン開始前にシーズン全体の需要予測を行う長期予測モデルと、シーズン中に日々のオペレーションで予測値の調整を行う短期予測モデルで構成されている。表 2 に、予測モデルに使用された目的変数と説明変数を示す。なお、「フラグ」は、ダミー変数として入力していることを示している。また、日付情報のうち「月」「日」「曜日」は順番を

表 2: 予測モデルの目的変数と説明変数

	説明変数	目的変数
長期予測モデル	-日付 (月/日/曜日/週番号)-出品数 -序盤/終盤フラグ -オフ率構成比 (0,10,20,...,90(%)) -気温 -気温過去 5 日分の分散 -前日気温差	販売数
短期予測モデル	-日付 (月/日) -気温実測値 -気温過去平均 -天候概要フラグ (晴_日中/雨_夜間,... など) -過去平均気温と実測値との差 -セール強度 -気温 5 日連続 t 度以上/以下フラグ (t = 0, 10, 20) -オフ率構成比 (0,10,20,...,90(%))	長期予測モデルの残差

考慮したいため数値データとして入力している。長期予測モデルの入力変数である出品数と気温については、学習時には実測値を用い、予測時にはシーズン前に設計した大まかな予測出品数と過去のデータから算出した平均気温を用いた。

長期予測モデルにより大まかなワンシーズンの販売数を予測し、短期予測モデルにより直近の新たに得た情報を用いて残差を予測し、長期予測モデルによる予測値と足し合わせることで販売数予測値を調整する。この 2 段階の手順で日別の販売数を予測する。実運用における役割として、長期予測モデルでは、計画段階で使える変数より長期的な日別販売数予測を行うことで、ピーク時期やその前後の動向の様子など対象期間の売れ行きの大まかな傾向を把握し、出品計画設計の手がかりとすることが期待できる。短期予測モデルでは、シーズン中のより多種で新しい情報を用いてより実際に近い販売数に予測値を調整でき、週周期やより具体的な直近の予測値を把握し、実際の出品数の策定・決定の支援が期待できる。結果として提案手法により、用いることのできる情報を無駄なく利用し定量的かつ的確に需要を把握した元で、在庫情報などを踏まえて「売れ残り商品を極力出さないように」などの戦略的な出品が可能となることが期待できる。また、表 2 に示した特徴量はモデルにより自動的に加味されるため、それらの影響を出品計画の際に改めて深く考慮する必要はなく、出品計画の負担軽減につながることを期待できる。

4. 実データ分析

本章では、提案モデルを実データに適用し得られた結果について分析し、前章で述べた実運用で期待できる役割を果たせるか否かについて、モデルの有用性を評価する。まず、長期予測モデルの基本的な精度を評価するために、最も回帰問題によく用いられている手法の 1 つである重回帰分析と、最新の時系列予測手法の Prophet により構築した予測モデルと精度を比較する。次に、短期予測モデルによる予測値の修正について評価するために、過去の一定期間について調整を行い予測精度の変化を観察する。

4.1. 分析条件

対象データは、買取日が 2015 年 12 月 1 日から 2020 年 11 月 15 日の春夏商品に関する販売履歴データである。本分析では、予測対象年から過去 2 年間の観測データを学習データとして、長期予測モデルを構築する。さらに、構築した長期予測モデルにより 2019 年度と 2020 年度について予測、実績データとの残差を算出し目的変数として用いて短期予測

表 3: 長期予測モデルの機械学習アルゴリズム別予測精度

	重回帰分析	Prophet	LightGBM
春夏-中	26.05	39.14	23.15
春夏-強	36.40	53.79	29.49
春夏-弱	53.89	58.60	43.35

表 4: 短期予測モデルによる予測精度の変化

	長期予測モデルのみの調整前	短期予測モデルによる調整後
春夏-中	27.14	17.12
春夏-強	22.27	10.74
春夏-弱	26.32	20.47

モデルを構築する。そして、2020 年度のシーズン最後の 2 週間の残差を短期予測モデルにより予測し、長期予測モデルの予測値に足し合わせることで調整する。なお、春夏商品のシーズン期間は、2 月から 8 月末とする。各予測モデルには LightGBM を用い、パラメータは Optuna [2] により最適化された値に設定した。ここで、社会情勢や社内トラブルなど、予期しないイベントの影響により明らかに予測不可能な販売数の変動は本研究の予測対象外とする。そこで、予測精度に対する評価指標として、外れ値の影響を受けにくい各データに対する Absolute percentage error(以下、APE)の中央値を用いる。なお、 n 番目のデータに対する APE は以下の式 (1) で算出する。ただし、 N 個のデータにおける n 番目の実測値を y_n 、予測値を \hat{y}_n とする。

$$APE_n = \left| \frac{\hat{y}_n - y_n}{y_n} \right| \quad (1)$$

過去の運用実績を鑑み、実運用への適用に対して十分に信頼性があると判断する閾値を、APE の中央値 30%以下と設定した。

4.2. 結果と考察

表 3 は、複数の機械学習アルゴリズムを用いて予測モデルを構築した場合の 2020 年度のシーズン全体の予測精度である。また、表 4 は、シーズン最後の 2 週間に関して、長期予測モデルの予測値による予測精度と、短期予測モデルにより修正した後の予測精度である。表 3 より、LightGBM により構築されたモデルが最良の精度で、「春夏-中/強」では十分に信頼性があると判定できる。よって長期予測モデルは、的確に傾向を把握し、シーズン前の大まかな出品計画の支援となる役割を果たすことが期待できる。また表 4 より、短期予測モデルの修正により全カテゴリで精度が向上していることがわかる。よって短期予測モデルは、より豊富な情報を用いてより実績に近い方向へ予測値を調整する役割を果たすことが期待できる。以上より、各モデルの実運用における有用性が示された。

5. ビジネスでの実運用による実証実験

本章では、提案手法によって得られた販売数予測値を、ZO-ZOUSED で新規出品される商品に対する出品計画決定の際の新たな指標として利用する実証実験を行う。また、販売実績の分析と考察を行うことで、実運用における提案手法の有

表 5: 売れ残り率の変化

傾向スコアによるマッチング	変化幅 (%)
無	-2.3
有	-3.1

効性を検証する。

5.1. 分析条件

ZOZOUSED の一部の春夏商品を対象に 2021 年 2 月から 8 月末に渡り以下のように実施する。まずシーズン前に、出品管理者に長期予測モデルの予測によりシーズン全体の需要を大まかに把握し、2 月からの出品開始の妥当性や在庫が十分あるかなど確認してもらう。シーズン中は、短期予測モデルを最新のデータを用いて学習し直近の需要を予測する。この予測値をベースとして、出品管理者にその先 1 週間で実行する出品数を決定してもらう。これを週次で繰り返し行い、8 月末に得られた実績について精度と有効性を評価する。精度評価は、式 (1) で表される APE の中央値により行うが、実務上は 30% 以下であることが望ましいとされている。有効性評価は、出品後 60 日連続で掲載されていても売れない商品を売れ残り商品とした場合の、出品数のうち売れ残り商品の占める割合と定義した売れ残り率により行う。過去最も平均的な出品をした 2019 年度の実績と比較し、減少していれば提案モデルが有効であると言える。ただし、気象情報や EC サイトの様子など、背景要因の膨大な組み合わせにより比較対象の 2019 年度と提案手法を実運用した 2021 年度が環境的に異なることも考えられる。すなわち、単純に全実績を集計し得られた売れ残り率の変化は、提案手法の効果ではなく環境の違いによるものである懸念がある。そこで、傾向スコアを用いて、全実績から環境をできる限り揃えるようマッチングした元でも、売れ残り率を比較検証する。

5.2. 結果と考察

予測精度を各週で計算した結果を図 2 に示す。なお、長期予測モデルによる調整前の結果を黒色点線、短期予測モデルによる調整後の結果を黒色実線、目標値を灰色実線で示している。また、右の数字は全体で計算した APE の中央値である。さらに、春夏商品の各月の売れ残り率を 2019 年度と 2021 年度で算出し比較した結果を、2019 年度 3 月の売れ残り率を 1 とした場合の指数として図 3 に示す。加えて、シーズン全体での売れ残り率の減少幅を計算した結果を傾向スコアによるマッチングの有無別に表 5 に示す。

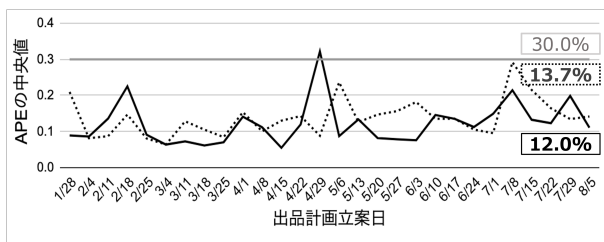


図 2: 各週の精度評価値

図 2 より、目標値より低い値をシーズン通しておおよそ保っており、実ビジネスにおいても十分な精度を持続するこ

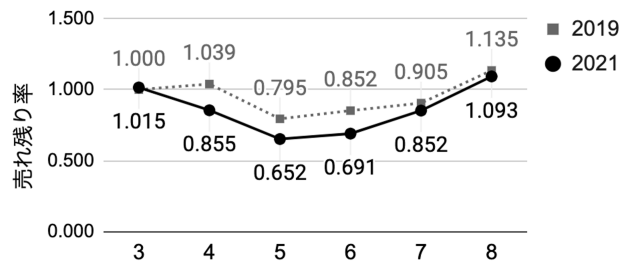


図 3: 月別売れ残り率

とがわかる。また図 3 より、3 月を除く全ての月で売れ残り率は減少し、全体でも 2.3% 減少している。さらに、傾向スコアによるマッチングを行なった元では 3.1% の減少と評価された。よって、実ビジネスにおいても売れ残り率減少に貢献し、出品計画の支援に有効であると言える。以上より、提案手法の信頼性と有効性が示された。以上の結果より、提案手法は、研究目的である需要に沿った戦略的な季節性のある中古アパレル商品の出品を可能にすると考えられる。

6. まとめと今後の課題

本研究では、中古ファッション EC サイトの季節商品を対象とし、出品計画において需要と供給のバランスが取れた最適な意思決定を実現する、実運用を想定した機械学習ベースの需要予測モデルから成る手法の提案を行なった。具体的には、シーズン開始前にシーズン全体の需要予測を行う長期予測と、シーズン中に新たに得た情報を用いた残差予測による予測値の調整を行う短期予測の、2 つの販売数予測モデルから成る 2 段階構成の手法を提案した。さらに、過去の販売履歴データに適用し、分析結果を通じて提案手法の有用性を示した。加えて、提案の予測値を、新規出品のための新たな指標として加えた元で決定した出品数を実行する実証実験を設計・実施することで、実証的に提案手法の信頼性と有効性が示された。また、効果検証の観点からも有効性を確認した。以上より、当 EC サイトにおいて、出品計画立案の支援となる有用なモデルを構築することができた。

今後の課題としては、春夏商品より高価格帯で販売数が多く、異なる傾向を見せる秋冬商品を対象とした実証実験の実施である。本研究の結果を踏まえてモデルおよび実証実験の設計を改良し、2021 年 8 月中旬より運営を行なっている。これらの結果についてまとめ、より詳細な考察・手法の改良を行うことが挙げられる。

参考文献

- [1] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, Vol. 30, pp. 3146–3154, 2017.
- [2] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2623–2631, 2019.