

リアルタイム人口統計情報を用いた店舗需要予測における 高需要帯の精度改善

篠田謙司^{†1} 山田将人^{†1} 高梨元樹^{†2} 長谷川大輔^{†2}
坪井哲也^{†2} 深澤佑介^{†1} 木本勝敏^{†1}

概要：飲食店において売上等の需要を予測することは、スタッフのシフトスケジューリングや店舗の事前準備を行う上で重要なタスクである。特に、イベント開催や天候変動を契機として通常より高い売上が発生する高需要状態を事前に予測することが求められている。一方で、従来より過去の売上データなどに基づいて未来の売上を推定する手法は多く提案されているが、これらの手法は全体精度を向上させることを目的としており、店舗運営上特に重要となる高需要状態での精度を重視した検討・評価は行なわれていない。本研究では、過去の売上データと天候データ、およびリアルタイムの人口統計データを用い、全体精度を大きく損なわずに高需要状態での予測精度を改善する手法を提案する。提案手法を飲食店における実売上データに適用し、予測精度を評価することで、その有効性を示す。

キーワード：飲食店、需要予測、異常検知、機械学習、リアルタイム人口統計データ

1. はじめに

飲食店において、食材調達やシフトスケジューリングなどの店舗運営の効率化を行う上で、的確な需要予測を行うことは重要なタスクである。各店舗を運営する店長は、過去の売上実績をもとに自ら予測を立て、その予測結果をもとに効率化を図ろうとするが、熟練度によって精度がバラつく。例えばベテラン店長は、店舗周辺の状況やイベントがどのように売上に影響を与えるかについて十分なノウハウを有しており、正確な予測をもとに店舗運営の効率化を行うことができる。一方で、新人店長にはそのような知見が少なく、結果として店舗運営を円滑に行うことができず、スタッフの離職やロイヤルティの低下につながる。

特に店舗周辺のイベント等を契機する高需要状態に関しては、ベテラン店長であっても的確に予測することは難しい。このような状態の予測がうまくいかなければ、想定外の需要増によって店内準備や指示出しに支障が生じ、料理の提供が遅れたり待ち時間が長くなるなど想定通りに店舗運営を行えないことになってしまう。

近年、このような需要予測の問題において、機械学習を用いた予測が盛んに行われている[1]。今回我々は、過去の店舗の売上データと、人々の統計的な位置情報等から、特に高需要状態に着目した売上予測を行う機械学習アルゴリズムを提案する。

本研究における貢献内容は次の3点である。

- 1) リアルタイムの情報（リアルタイム人口統計データおよび天候データ）を主な入力として、高需要状態を予測可能なモデルを構築したこと。
- 2) 通常の需要状態を予測可能なモデルと、高需要状態を予測可能なモデルの出力結果から、現在の状態が高需要状態か通常状態かを判別し、その状態に応じて予測結果をアンサンブルする手法を提案したこと。

- 3) 高需要帯での精度評価を行うための指標として、高売上時予測精度と高予測時予測精度を導入し、提案手法が高需要帯において精度改善することを示したこと

1.1 関連研究

これまで店舗の売上に関わる予測は数多く行われてきた。Liuらは、商品の売上実績データと天候データを用いて、商品の売上の度合いを予測する手法を提案している[2]。Yangらは発注の最適化を目的として、大量のPOSデータを分析し、直前のPOSデータの傾向を用いて直後数時間の売上予測を行う手法を提案している[3]。

しかし、これら既存研究による手法の多くは、全体精度の向上を目的とした特徴量設計やモデリングが中心となっており、イベント開催や天候変動を契機とした高需要状態での予測精度を重視した取り組みは行なわれていない。

一般に、全体精度の向上をめざした売上予測モデルは、売上変動の周期性が重視され、通常需要状態を得意とするモデルになりやすく、イベント開催や天候変動を契機とした高需要状態における予測精度は期待しにくい。一方で、高需要状態を重視したモデルを構築すると、全体精度を大きく劣化させてしまう傾向がある。

1.2 本研究の目的

本研究では、基本手法として構築された通常需要状態を得意とするモデルと、高需要状態を得意とするモデルとをそれぞれ構築し、さらに、これらのモデルを適切な重みで組み合わせるための需要帯分類モデルを構築することで、全体精度を大きく損なわずに、高需要状態での予測精度を改善するアンサンブル手法を提案する。

本稿では、2章で問題設定とリアルタイム人口統計デー

^{†1} 株式会社 NTT ドコモ
NTT DOCOMO, INC.

^{†2} 株式会社サイゼリヤ
Saizeriya Co., Ltd.

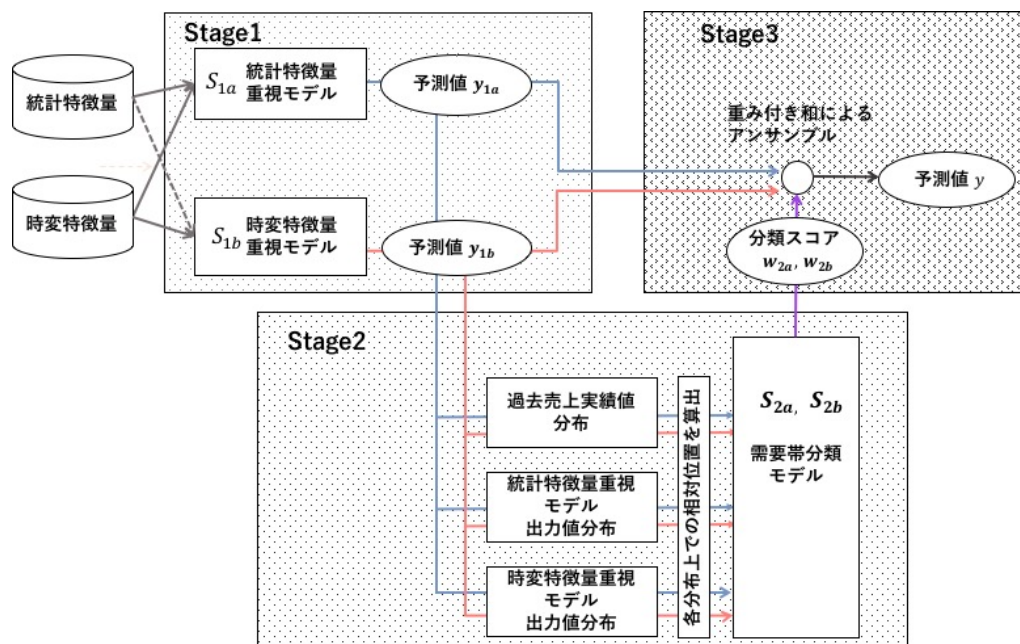


図 1. 提案モデルのアーキテクチャ

タの概要を述べ、3章で提案手法について詳説する。4章で高需要帯の精度を評価するための指標を定め、実データを用いて提案手法の評価を行う。

2. 問題設定

本研究では、店舗ごとに取得可能な POS データから作成される売上データと、店舗周辺のリアルタイム人口統計データ、および天候データを入力とし、1 時間先の店舗ごとの売上値を算出する回帰モデルを作成する。

2.1 リアルタイム人口統計データ概要

リアルタイム人口統計データは、携帯電話ネットワークの仕組みを利用して人口を推定した統計値である。モバイル端末と基地局の位置関係から、東京都内では 500m の空間粒度、10 分毎の時間粒度で生年代別での人口推定を可能としている。^{*a}

2.2 リアルタイム人口と売上の関係

飲食店における売上増と人口増減の関係を調べるために、イベントが多発する施設近辺の店舗において、ある期間で売上値と人口の変動をプロットしたものを図 2 に示す。赤枠で示す通り、イベント発生によって通常状態に対し人口が増大したときに、売上が通常時より増大していることが分かる。この結果は、リアルタイムな人口統計データを用

いることで高需要帯の精度向上を図れる可能性を示唆している。

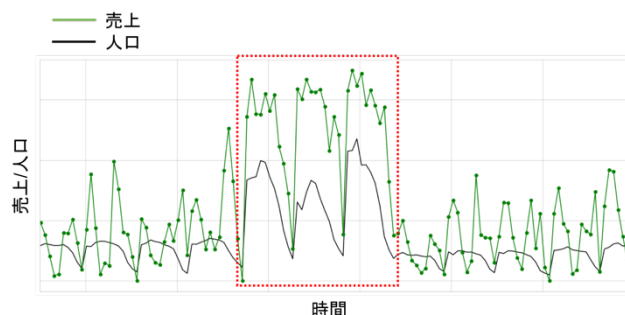


図 2. 1 時間ごとの売上値と人口の時系列グラフ

3. 提案手法

3.1 アプローチ

イベントを契機とするような高需要状態を予測するため、通常需要状態を得意とするモデルの他に、リアルタイムに変化する時変特徴量（リアルタイム人口統計データと天候データ等）を重視させることで高需要状態での予測を得意とする時変特徴量重視モデルを構築する。それらの回帰値を用いて、予測対象時の店舗需要状態が通常需要状態と高需要状態のいずれになるかを予測する需要帯分類モデルを構築し、その出力結果（分類スコア）に基づいて、回帰値それぞれに重みを持たせ組み合わせることで、全体精

a 本実験で使用する人口統計は、エリア毎や属性毎の集団の人数を示す情報であり、お客様個人を特定できる情報を一切含みません。したがって、この人口統計によりお客様の行動が他人に知られることはありません。なお、本実験で使用する人口統計は、モバイル空間統計ガイドラインを遵守

しております。

モバイル空間統計ガイドライン

https://www.nttdocomo.co.jp/corporate/disclosure/mobile_spatial_statistics/guideline/index.html

度を大きく損なわずに、高需要状態での予測精度の改善を実現する。

提案手法のアーキテクチャを図 1 に示す。各モデルには弱学習器のアンサンブル法の一つである LightGBM[4]を用いている。

3.2 特徴量の生成

3.2.1 空間粒度/時間粒度の変換

売上を予測するにあたり、我々はリアルタイムな人口統計データや天候データといった異種データを組み合わせてモデル構築を行う。各データはそれぞれ異なる空間粒度、時間粒度を持っていたため、それらを一律に揃える必要がある。今回我々は時間粒度を 1 時間、空間粒度は 4 次メッシュの定義のもとでの 500m グリッドとし、それに揃えるように各データの集約を行った。4 次メッシュとは国土地理院が定める 500m 四方のグリッドを指す。

売上データは、POS システムに登録された注文日時と各商品の金額から、1 時間ごとの売上合計金額を算出した。また各店舗の緯度経度情報を 4 次メッシュに変換した。リアルタイム人口統計データは、空間粒度は 4 次メッシュとして与えられているため、時間粒度のみ 20 分毎のデータへ集約を行なった上で、各時刻の最初の 20 分をその 1 時間での人口値とした。天候データは、1 時間ごとの予測値を作成した上で、空間粒度を 500m グリッドに変換した。

3.2.2 統計特徴量

店舗の売上値は季節や曜日（平日・休日・休前日・休後日等）、時刻によって変化する。このとき、ある店舗ある曜日のある時間台における平均値や標準偏差値といった統計情報は、その条件下における尤もらしい推定値を算出する際に重要な情報となる。今回我々は 1 ヶ月前の同店舗同曜日同時間台というように、特定の条件を課した上で売上値の統計量（平均値・標準偏差値・最大値・最小値）を計算し、これを統計特徴量とした。具体的な条件の例を表 1 に示す。

表 1. 統計特徴量作成のための条件

カテゴリ	例
曜日	休日, 平日, 休前日, 休前日
気象	気温レベル, 雨量レベル
時間	同時間台, 1 ヶ月前, 1 年前

3.2.3 時変特徴量

通常時の売上に対する乖離を示す高需要状態を追うには、店舗周辺の人口変動や各時間台での気象状態を考慮することが重要である。特にリアルタイムな人口統計データは、周辺のイベント発生時には通常と異なる動きをする

と考えられ、高需要状態を予測する上で重要な因子となりうる。

我々は店舗が属する 4 次メッシュと、その周辺 8 メッシュを対象に、その時変性を時間差分・空間差分をとることで特徴量化した。模式図を図 3 に示す。

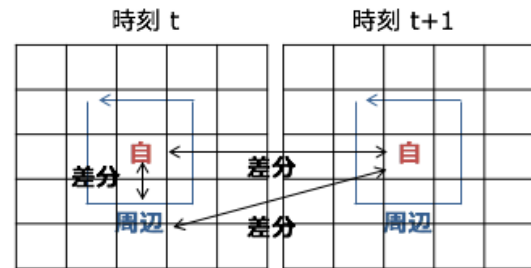


図 3. 人口データの時変特徴量

3.2.4 時刻情報の変換

現在の時刻情報として、曜日情報や時間台情報を用いる。曜日情報は当日の平日・休日・休前日判定に基づいた 2 値データで表現した。11 時台、12 時台といった時間台情報は下記式に基づいて周期情報への変換を行なった。

$$\alpha = \sin \frac{\pi}{12} t, \quad \beta = \cos \frac{\pi}{12} t, \quad \{0 \leq t < 24\} \quad (1)$$

3.2.5 目的変数の正規化

人口統計データや天候データを学習に用いる上で、予測対象の売上が通常時と比べてどの程度変容しているかを定量的に表現することは重要である。3.2.2 において統計特徴量を作成した際と同じように、同店舗同曜日同時間台のような条件を課して分布を計算し、その平均と標準偏差を用いて標準化を行なった。変換式を式 2 に示す。ここで X は条件付けられる特徴量の集合、 x はその条件、 Y は正解となる売上値の集合で、 $Y_{X=x}$ とは $X=x$ と条件付けられるときの Y の集合を意味する。また期待値と分散を E, V と表している。

$$Y'_{X=x} = \frac{Y_{X=x} - E(Y|X=x)}{\sqrt{V(Y|X=x)}} \quad (2)$$

3.3 Stage1:統計特徴量重視モデル/時変特徴量重視モデル

図 1 で示す統計特徴量重視モデル S_{1a} とは、3.2.2 のように作成した統計特徴量と、3.2.3 のように作成した時変特徴量の全てを用いて学習を行なったモデルである。また時変特徴量重視モデル S_{1b} は、 S_{1a} で用いた特徴量のうち、一部の統計特徴量と、全ての時変特徴量を用いて学習を行なったモデルである。

これら 2 つのモデルそれぞれから、売上予測値としての回帰値が出力される。

3.4 Stage2:高需要帯分類モデル/低需要帯分類モデル

Stage2 では予測対象の需要値が高需要帯に属するかそうでないかを判別する高需要帯分類モデル S_{2a} と、低需要帯に属するかそうでないかを判別する低需要帯分類モデル S_{2b} を構築する。説明変数には、Stage1 の回帰によって得られた統計特徴量重視モデルの出力値がなす分布、時変特徴量重視モデルの出力値がなす分布、および売上データがなす分布のそれぞれに対して、各予測値がとる相対関係を表現する量を用いる。

図4に分布と予測値の関係の模式図を示す。まず、同店舗同曜日同時間台と条件を課して、あらかじめ、 S_{1a} 、 S_{1b} のそれぞれが成す予測値分布および正解分布を求めておく。予測対象店舗・曜日・時間台における S_{1a} の予測値が、各分布において高需要状態にあるかどうかを判定することで、分布に対する予測値の相対関係を定量化することができる。これらの量は S_{1a} 、 S_{1b} は事前情報として持たない量であり、分類を行う上で有用な情報になると考えられる。

高需要状態と判定する閾値は、各曜日各時間ごとに、各分布の平均 m と標準偏差 σ を用いて $m \pm 1.2\sigma$ と定義する。このとき、目的変数は正例に対して負例が十分に多く、大きな偏りを持っているが、一般にこのようなデータで分類問題を解くと十分多い負例に寄った推論を行ってしまう。それを避けるために、負例に対してダウンサンプリングを行ない、割合を正例:負例=1:2とした。

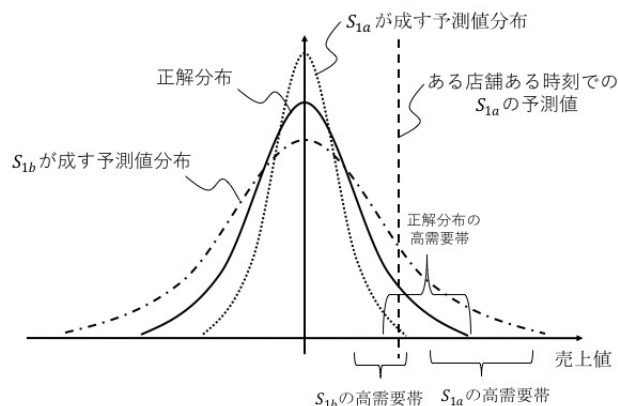


図4. 分布の相対関係を用いた定量化

3.5 Stage3:重みづけ線形和によるアンサンブル

Stage2 のモデルが出力する分類スコアを用いて Stage1 の回帰値のアンサンブルを行う。すなわち、 S_{1a} 、 S_{1b} の回帰値をそれぞれ y_{1a} 、 y_{1b} 、 S_{2a} 、 S_{2b} の分類スコア値をそれぞれ w_{2a} 、 w_{2b} として、予測値 y を下記の式に従い計算する。

$$y = \{[w_{2a} \max(y_{1a}, y_{1b}) + (1 - w_{2a}) \max(y_{1a}, y_{1b})] + [w_{2b} \min(y_{1a}, y_{1b}) + (1 - w_{2b}) \min(y_{1a}, y_{1b})]\} / 2 \quad (3)$$

また実験では比較のため、Stage2 の2つの回帰値を50%ずつ足し合わせる単純平均アンサンブルと、一方の回帰値の重みをもう一方の回帰値とする調和平均アンサンブルも計算する。調和平均に関する計算式を下記に示す。

$$y = \frac{2y_{1a}y_{1b}}{y_{1a} + y_{1b}} \quad (4)$$

4. 実験

4.1 データの詳細

本稿では表2に示すような売上データ、人口統計データ、天候データの3種類のデータに対して3.2の処理を行ってモデルの学習・検証を行った。店舗の売上データは、2週間毎に得られるPOSデータの集約情報を用いる。リアルタイムな人口統計データは2.1で述べた通りの10分毎に500m粒度で得られるデータを用いた。また天候データは、5kmグリッドで各時間から39時間後までの降水量・雲量・気温等を予測するMSMデータを用いた。

表2. データの詳細

対象地域		関東近郊
期間		2016年5月1日から 2018年10月31日まで
売上データ	店舗数	347店
	取得頻度	2週間毎
	空間粒度	店舗
	時間粒度	1時間
人口統計データ	取得頻度	10分毎
	空間粒度	500mグリッド
	時間粒度	10分
天候データ	取得頻度	毎日0時から3時間毎
	空間粒度	5kmグリッド
	時間粒度	1時間

4.2 学習・評価データの期間

提案手法では各Stageで異なる学習モデルを必要とする。今回は図5のように各モデルで期間を揃えて学習と評価を行なった。

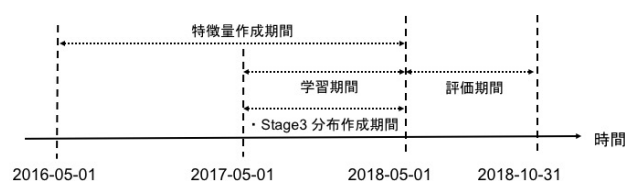


図5. 学習・評価の期間

4.3 評価指標

高需要帯の評価を行う際に、実際の売上が高売上のときに精度が良いのか、あるいは学習器が高予測を行なったときに精度が良いのか、それぞれ評価を行うことは予測の信頼性の観点から重要である。例えばモデルが高売上帯の予測を出し続けられれば、高売上時の精度はよくなり見かけ上高需要帯の精度が改善されたように思えるが、このような場合には高予測時に限った評価を行うと著しく精度が劣化する可能性がある。本稿では評価区分を、全体需要に対する評価、高売上時の評価、高予測時の評価の3区分で行う。

提案手法の有用性を評価するため、上記3つの種別に関して、各モデルの絶対平均誤差 (MAE) の S_{1a} に対する改善率 (IR) で評価を行った。例として、 S_3 の高売上時 MAE の定義を式5に、改善率 IR の定義を式6に示す。ここで $\hat{y}_i$ は各店舗の実際の売上値、 y_i は予測した売上値、 n は店舗数である。また P はモデルが高売上と予測した回帰値が成す集合である。

$$MAE_P^{S_3} = \frac{1}{\#P} \sum_{i \in P} |\hat{y}_i - y_i| \quad (5)$$

$$IR_P^{S_3} = \frac{MAE_P^{S_{1a}} - MAE_P^{S_3}}{MAE_P^{S_{1a}}} \times 100\% \quad (6)$$

4.4 実験結果

説明変数やアンサンブル方法を変更した各モデルに対して評価した結果を表3に示す。ここで表の統計特徴量重視モデルである S_{1a} を基本手法として、提案手法の改善率を評価する。調和平均、単純平均、提案手法はそれぞれ3.5で述べたアンサンブル法を適用した結果である。

表3より、全体精度において各アンサンブル手法は統計特徴量重視モデルから大きく劣化はしていないことが分かる。一方で高需要帯に関しては、特に提案手法において他のアンサンブルに比べてより高売上時、高予測時ともに精度改善していることが分かる。これは、Stage2 で出力された分類スコアを重みとして用いることで、高需要帯となりうるかそうでないかの分類スコアに基づき選択的にアンサンブルできていることを示している。

表3. 異なる入力データによるモデル評価結果

モデル	全体需要 改善率	高売上時 改善率	高予測時 改善率
S_{1a} 統計特徴量重視	-	-	-
S_{2a} 時変特徴量重視	-1.1%	+0.73%	+2.5%
単純平均法	-0.15%	+0.46%	+2.1%
調和平均法	-0.13%	+0.28%	-2.1%
提案手法	-0.13%	+1.45%	+2.4%

次に、提案手法のアンサンブルにおいて、損失関数を変更した場合の評価を行う。ここでは表3の結果を得る際に用いた損失関数としての RMSE の他に、我々が定義する weighted RMSE を損失関数として用いた場合の結果も示す。この損失関数は二乗誤差の因子に正解値の絶対値をかけ、異常な売上に対しより鋭敏に反応することを意図するコスト・センシティブな関数である[5]。式7に weighted RMSE の数式を、図5に指数 a に対する weighted RMSE の振る舞いを示す。ここで a, c は共に定数とする。

$$L = (\hat{y} - y)^2 (|y|^a + c) \quad (7)$$

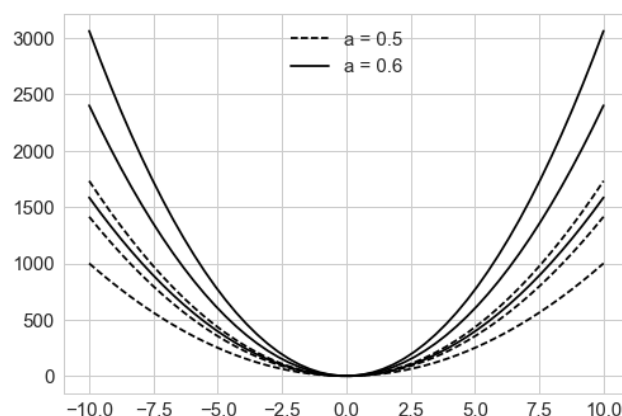


図5. Weighted RMSE の振る舞い

評価結果を表4に示す。通常の RMSE を損失関数とする場合に比べて、全体需要や高予測時の精度は劣化してしまうが、高売上時の精度は大きく改善することが分かる。これは weighted RMSE の持つ指数 a を変更することで、高売上時の予測精度をより重視するかどうかを調整できることを示している。

表4. Weighted RMSE を用いた場合の評価結果

損失関数	全体需要 改善率	高売上時 改善率	高予測時 改善率
RMSE	-0.13%	+1.45%	+2.4%
weighted RMSE ($a = 0.5, c = 0$)	-2.2%	+10.1%	-0.2%

ある店舗のある期間における、売上の正解値と S_{1a} と S_3 の予測値を図6に示す。正解値を緑の実線、 S_{1a} の予測を赤の破線で示し、 S_3 の予測を青のエリアプロットで示している。提案手法である S_3 が、 S_{1a} に対して高需要状態をより追跡していることが定性的にわかる。

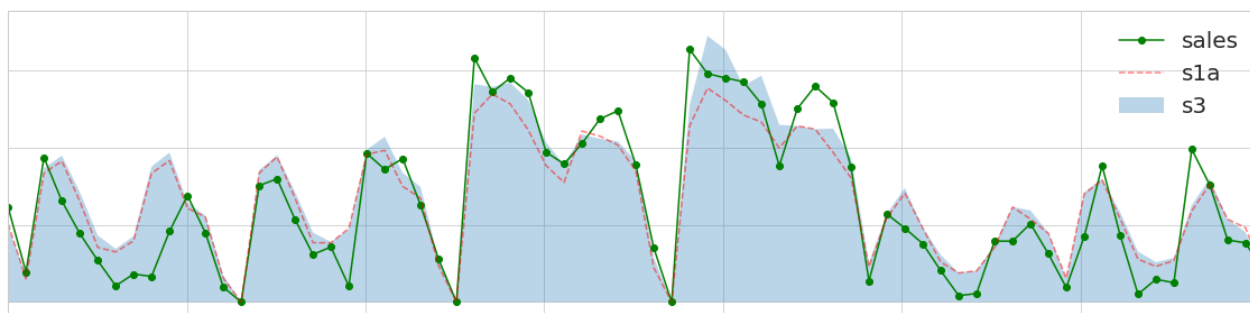


図 6. 予測の時系列グラフ

5. まとめ

本稿では、通常需要状態での予測を得意とする統計特徴量重視モデルに対して、時変特徴量を重視させることで高需要状態での予測を得意とする時変特徴量重視モデルを構築した。さらに、予測対象時の店舗需要状態が通常需要状態と高需要状態のいずれになるかを予測する需要帯分類モデルを構築し、その分類スコアを用いて重み付けアンサンブルを行うことで、全体精度を大きく損なわずに、高需要状態での予測精度を改善する手法を提案した。

また高需要帯における評価指標を導入した上で、飲食店の実売上データを用いた評価を行い、損失関数に RMSE を選択した場合に、既存手法に対して高売上時の予測精度が 1.45%、高予測時の売上精度が 2.4%改善することを示した。さらに、損失関数に weighed RMSE のようなコスト・センシティブな関数を採用することで、全体精度と高売上時の精度を調整できることを示した。

今後は、アーキテクチャ内の各個別モデルに用いている LightGBM の置き換えとして、時系列回帰には LSTM[6]、需要帯分類には GAN[7,8]や Autoencoder[9,10]等の深層学習モデルをそれぞれ適用し、さらなる精度向上に取り組んでいく。

謝辞

本研究を実施するにあたり、共に活発な議論をして頂いたチームメンバーの皆様に謹んで感謝の意を表する。

参考文献

- [1] Agnieszka Lasek, et al., Restaurant Sales and Customer Demand Forecasting: Literature Survey and Categorization of Methods, International Summit, Smart City 360° (2016)
- [2] Xin Liu, et al., Food Sales Prediction with Meteorological Data - A Case Study of a Japanese Chain Supermarket, Data Mining and Big Data (2017)
- [3] Chao-Lung Yang, et al., Short-Term Sales Forecast of Perishable Goods for Franchise Business, IEEE International Conference on Knowledge and Smart Technology (2018)
- [4] Guolin Ke, et al., LightGBM: A Highly Efficient Gradient Boosting Decision Tree, NIPS (2017)
- [5] Flavian Vasile, et al., Cost-sensitive Learning for Utility Optimization in Online Advertising Auctions, KDD (2017)
- [6] Sepp Hochreiter, et al., Long short-term memory, Neural Computation (1997)
- [7] Ian J. Goodfellow, et al., Generative Adversarial Nets, NIPS(2014)
- [8] H. Zenati, et al., Efficient GAN-Based Anomaly Detection, ICLR (2018)
- [9] Geoffrey E. Hinton, et. al, Reducing the Dimensionality of Data with Neural Networks. Science (2016)
- [10] Chong Zhou, et al., Anomaly Detection with Robust Deep Autoencoders, KDD (2017)