

ラベル伝搬法に基づく店舗でのキャンペーン推薦方式

Campaign Recommendation Method for Retail Sales by Label Propagation Algorithm

宮田 淳司* 伊藤 智祥*
Atsushi Miyata Tomoaki Itoh

ラベル伝搬法を店舗データに適用し、小売店舗向けにキャンペーン推薦を行う方式について解説する。従来、キャンペーン未実施の店舗でのキャンペーン効果の予測は困難であった。本稿ではPOS・商圈からなる店舗データの類似性を考慮して、キャンペーン実施店の結果から効果予測する方法を提案する。評価実験では、本方式はキャンペーン効果の予測精度が向上し、売上額にして12%の向上が期待できると試算され、有効性が確認できた。

We explain the function of campaign recommendation method for retail stores. In our method, we apply the label propagation algorithm, one of the semi-supervised learning algorithms, to retail sales data. We address the difficulty of campaign recommendation for stores not running campaigns. Our proposal method use similarity of store data based on POS and trade area data. We show the validity of our method by that experimental results show that the prediction accuracy increase and amount of campaign sales increase is estimated to be 12% by our method.

1. 店舗分析AIと小売りでのキャンペーン推薦

現在の小売業界では、店舗改善は主にID-POS (Point Of Sale) データを用いた顧客購買行動の分析、あるいは店長の経験や勘によるところが大きい。しかし店舗内外での顧客の購買促進のためには、購買に至るまでの顧客の行動をより詳細に把握する必要がある。それらの知見はID-POSデータに加えて、地域性や店内行動、ソーシャルメディアなどのデータの分析を行うことで得られる。しかし扱うデータの種類が増えると、人手による分析でデータ間の関係性を見つけるのは困難となる。

そこで筆者らは、ID-POS、店舗内動線など多様なデータから店舗での分析・提案を自動化する店舗分析AIシステムの開発を行っている。このシステムは人手で発見するのが困難な、膨大なデータの規則性や関係性をAIにより見だし、店舗改善のための提案を生成する。

店舗分析AIの提案の1つとして、複数の候補から店舗に合わせたキャンペーンを推薦するキャンペーン推薦機能の研究開発を行っている。キャンペーンの候補はチラシ、クーポン、店内サイネージなど多数あるが、店舗では1つのキャンペーンの実施にも多くの手間と時間がかかる。本機能はこのような問題を解決するもので、複数の店舗での業務データ (ID-POS・商圈・店舗内動線データなど) や実施効果のデータ (キャンペーン前後での売り上げ変化データなど) を収集する。これらのデータに基づきキャンペーン未実施の店舗で有効なキャンペーンを推薦する。

本稿では、半教師あり学習すなわち、正解データが一部欠損してもモデル学習が可能な機械学習手法である、ラベル伝搬法[1]を用いたキャンペーン推薦の方式を提案

する。従来の方式との比較は第1表のとおりである。従来方式には、購買データや商圈データから実験店舗を選定し、実験店舗での効果からクロス分析を行ってキャンペーン推薦を行う方式、店舗の各データでの線形回帰モデルから予測を作成する方式、交通網を数理モデルのグラフに変換し、接続特性で各地域を分類してキャンペーンを推薦する方式などが提案されている。

第1表 キャンペーン推薦方式ベンチマーク

Table 1 Campaign recommendation methods benchmark

	ラベル伝搬法 (提案方式)	クロス分析	線形回帰 モデル	グラフ構造 分析
キャンペーン未実施店での推薦	○	×	×	×
推薦精度	○	○	○	△
必要なデータの収集しやすさ	○	×	○	○
特徴	町村間の類似性活用により精度向上	綿密な実験計画の策定	多数データで回帰モデルを作成	交通網の特性から町村をクラスタ分類

2. ラベル伝搬法でのキャンペーン推薦方式

2.1 提案方式におけるキャンペーン推薦の流れ

キャンペーン推薦は各候補キャンペーンの効果を予測し、より良い効果が期待できるものを薦める。キャンペーン効果の大小は、店舗商圈のなかでキャンペーンに好反応する顧客の多寡に強く影響を受ける。好反応を示す顧客の数は、地域の人口・世帯数・来店客数などの地域の購買者属性分布が類似している地域では、同等になると考えられる。以上の考えからキャンペーン推薦対象店

* 要素技術開発センター

Core Element Technology Development Center

舗の効果予測は、店舗商圏内の各町村の効果ランクを集約することで行う。町村のキャンペーン効果ランクを予測するため、過去のキャンペーン実績に基づいて付与した町村の効果ランクと、その町村の購買者属性分布を用いてモデル化する。具体的には、効果ランクとは、キャンペーン前後での売り上げの変化額(=キャンペーン後の売り上げ-キャンペーン前の売り上げ)に応じて、売り上げの上昇額が高い順にA～Dの4段階のランク付けを行ったものである。

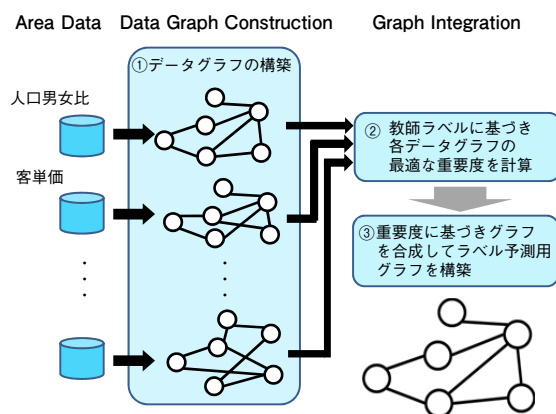
2.2 ラベル伝搬法によるキャンペーン効果予測

ラベル伝搬法では「グラフ」と呼ばれる、町村間の類似関係をリンクで表す数理モデルを用いる。ノードは町村を表し、1つのグラフには全店舗の商圏内の町村が含まれる。各ノードは町村データとラベルの2種類の情報を保有している。町村データとは、例えばID-POSから取得可能な町の人口の男女比や客単価である。1つの町村は複数の店舗商圏にまたがることもあり、その場合店舗ごとにキャンペーン効果が異なることがあるため、地理的に1つの町村でも、グラフ内では店舗ごとに別のノードとして扱う。

町村間のリンクの強度は町村データ同士の類似性、例えば人口男女比の類似性によって決定し、町村データ間の類似性が強いほどリンク強度は高くなる。ラベル伝搬法ではリンク強度が強い町村同士のキャンペーン効果が類似するという仮定のもと、リンク強度に従ってキャンペーン未実施の店舗商圏の町村のノードに対する効果ランクラベルを予測する。

キャンペーン効果を予測するグラフのリンク作成方法は、第1図に示すとおり[2]、町村データ種別ごとに作成したデータグラフを合成することで、ラベル予測用のグラフを作成する。まず人口の男女比や客単価など町村データの種別ごとに個別にグラフを構築する(第1図①)。このときノードがもつラベルは各グラフで共通とする。各グラフの重要度はキャンペーンを実施済みの各町村のラベルを予測するときの誤差が最も小さくなるように決定する(第1図②)。計算した重要度に基づき各データのグラフを1つのグラフに合成して効果予測グラフを構築する(第1図③)。

ラベル伝搬法を含む半教師あり学習の特徴は、キャンペーン効果が判明している町村が例え少数であっても、町村データの類似性を用いて効果を予測することができることである。提案方式ではそれに加えてグラフの再構築により予測精度の向上を図った。ラベルを予測した町村のうち、キャンペーン実施済みの町村(ラベル付与済みの町村)と予測グラフ上で類似性が高い町村の予測ラ



第1図 複数データを統合したラベル予測用グラフの構築

Fig. 1 Formulation of campaign label prediction graph by integrating multimodal data graphs

ベルから順に利用して、各町村データの種別ごとのグラフの重要度を再計算する。そして再計算した重要度でラベル予測用のグラフの再構築を行った。この再構築では、予測したラベルが正しいラベルと異なる場合、予測誤差が拡大するという欠点が存在する。そのため予測精度の向上には、変数の選別やリンクの設計など初期のグラフ設計が非常に重要となる。

3. 店舗向けキャンペーン推薦方式の評価

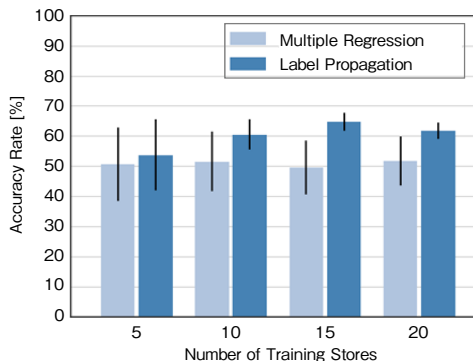
提案方式の精度評価のため、初夏の新CM放映・晩夏の新商品発売の2つのキャンペーンを利用して実験を行った。効果予測グラフは、2016年5月～6月の飲料売り上げの変化額に基づく町村のランクと2016年5月の町村データ、2016年7月～8月の飲料売り上げの変化額に基づく町村のランクと2016年7月の町村データを使用して、それぞれ作成した。町村データには町村別の人口・人口男女比・世帯数・昼間人口・有効会員数・店舗利用者数・店舗総売り上げ・店舗のべ利用者数・飲料利用者数・飲料総売り上げ・飲料のべ利用者数・店舗客単価・店舗月間単価・店舗までの距離・平均世帯収入のデータが含まれている。

実験では2つのキャンペーンのうちどちらが効果があるか、もしくは同程度に有効であるかの3つの選択肢に関して予測を行い、実際の結果と比較した。予測精度の指標としてキャンペーン推薦の正解率(=予測が実績と一致した店舗数÷予測対象の全店舗数)を用いる。実験では、町村データが判明している全163店舗のうち、少数の店舗を選んでキャンペーン実施店すなわち教師店舗とした。残りすべての店舗は効果不明としてこれを予測することで、正解率の算出を行った。教師店舗の選択は、店舗商圏内の町村の4つのランクが、同数程度となるよう選んだ。

第2図はキャンペーン実施済みの教師店舗を5店舗～20店舗と5店舗ずつ増加させた際の線形回帰モデルとラベル伝搬法での推薦精度を比較したグラフである。線形回帰モデルでは教師店舗でのキャンペーン前後の売り上げ変化額を学習し、町村データから残りの店舗商圏の各町村のキャンペーン効果を予測し、予測された値から各町村のランク分類を行った。第2図の結果より教師店舗数が5店舗・10店舗の場合には推薦精度の差は比較的小さく、学習に利用する教師ラベルが不足していることがわかる。しかし15店舗・20店舗と教師店舗数が増加すると線形回帰モデルでの推薦精度がおおよそ50%強程度でほぼ横ばいであるのに対し、ラベル伝搬法でのキャンペーン推薦精度は65%程度に増加しており、ラベル伝搬法によるキャンペーン推薦が精度が高いことがわかる。またラベル伝搬法での正解率のばらつきが小さくなることから、キャンペーンをどの店舗で実施したかによらず、効果予測精度が安定することがわかる。これらの結果はキャンペーン実績のデータ量が少なく、線形回帰モデルなどでは予測精度が上がらない状況において、ラベル伝搬法でのキャンペーン推薦が有効であることを示している。

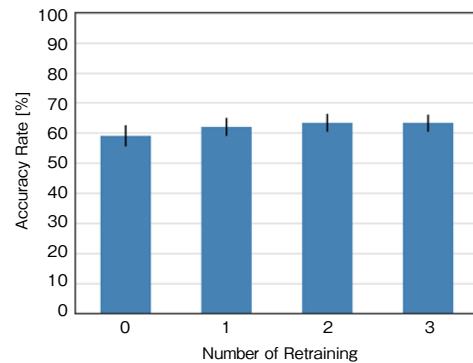
また予測したラベルを利用してグラフの再構築を行うことによる効果を確認する実験を行った。教師店舗数20店舗での結果を第3図に示す。この図からグラフ再構築を2回以上行うことで、再構築を行わない場合と比較して精度が5%程度向上することがわかった。以上からラベル伝搬法を活用したキャンペーン推薦において、予測したラベルを利用してグラフの再構築を行うことで推薦精度が向上することがわかった。

さらに、キャンペーン推薦の効果額の試算を行った。2つのキャンペーンの一方をランダムに選択して実施した場合と比較して、本方式によって推薦されたキャンペー



第2図 教師店舗数に対するラベル伝搬法と線形回帰モデルでのキャンペーン推薦の正解率

Fig. 2 Accuracy rate of campaign recommendation by label propagation and multiple regression analysis



第3図 効果予測グラフを再構築した際のキャンペーン推薦精度の変化

Fig. 3 Accuracy rate of campaign recommendation by number of retraining effect prediction graph

ンを実施した場合、キャンペーンによる売上額は、提案方式の方が12%向上すると試算された。

以上本提案方式を用いることで、POSデータと商圏データを用いて、過去に同様のキャンペーンを実施したことがなくても、複数キャンペーンの効果を予測し、予測結果に基づいてキャンペーンを推薦することが有効であることを示した。

4. 今後の取り組みと展望

本稿では一部店舗のキャンペーン実績に基づいて、他の店舗でキャンペーン推薦を行う方式を提案した。しかし本手法はキャンペーン推薦に留（とど）まらず、商品レコメンドに適用することも可能である。例えば町村の代わりに顧客間のグラフを作成する。顧客ごとのID-POSの購買やBeaconから棚への立ち寄り情報を取得して顧客間のリンクを作成し、顧客の商品選好ラベルを伝搬することで実現できる。今後は顧客行動データを活用して店舗改善を支援するさまざまな新たな機能の開発をさらに進め、小売店向けB2Bソリューションビジネスで貢献していきたい。

最後に研究を進めるにあたり、ID-POS・商圏データをご提供いただいた（株）トライアルカンパニーに深謝の意を表する。

参考文献

- [1] D. Zhou et al., "Learning with local and global consistency," NIPS 2003 Proceedings of the 16th International Conference on Neural Information Processing Systems, pp.321-328, 2003.
- [2] T. Kato et al., "Robust label propagation on multiple networks," IEEE Trans. on Neural Networks, vol.20, no.1, pp.35-44, 2009.