

製品別アソシエーション分析を用いた 季節関連性製品の抽出

千葉工業大学 社会システム科学研究科 修士課程

千葉工業大学 社会システム科学研究科 准教授

千葉工業大学 社会システム科学研究科 教授

河西 裕次郎

武田 善行

加藤 和彦

要約

ビッグデータを用いたデータマイニングでは、顧客の詳細なニーズを正確に把握するために、顧客関係管理（CRM）が重要な役割を果たしている。CRMの実践においては主にRFM分析を用いて、顧客満足度の向上を図り、企業の収益増加に繋げることを目的としている。また、アソシエーション分析を用いることで、有用な製品の組み合わせを明らかにし、マーケティング施策に反映させている。さらに、クラスタ分析を組み合わせることで、アソシエーションルールに基づいた顧客のセグメント化が可能となり、各クラスタの特性や行動パターンを詳細に理解することができる。

本研究では、RFM分析で顧客を分類し、次にアソシエーション分析を実施して四半期毎の購買傾向を詳細に解析した。その後、アソシエーションルールを基にクラスタ分析を行い、各クラスタの特徴的な購買パターンを抽出した。四半期毎のアソシエーションルールの違いを明らかにすることで、時期によって変化する製品の需要を把握し、顧客の詳細なニーズを予測することが可能となった。

キーワード

顧客関係管理, RFM分析, アソシエーションルール, クラスタ分析

1. はじめに

ビッグデータを用いたデータマイニングでは、顧客の詳細なニーズを推定するために、顧客関係管理（CRM）が盛んに行われている。CRMでは主にRFM分析を用いて顧客を分類し、各分類に対して適切な施策を実施することで、顧客満足度の向上を通じた収益増加を目的としている。五十嵐・松井・南・水田（2015）は、共同購入型クーポンサイトの購買データを用いて、顧客の購買傾向に関する有用な知見を得る方法論を示した。さらに、顧客の属性と購入ジャンルとの関係性について詳細に明らかにした。

アソシエーション分析の歴史は長く、百貨店のPoint of Salesデータの分析により、膨大なデータから意味のある関連性を抽出し戦略を練るために開発され、現在も広く使用されている。Agrawal and Srikant (1994) は、大規模なトランザクションデータベースからアイテム間の関連ルールを発見する問題を解決するために2つのアルゴリズムを提案した。さらに、これら2つのアルゴリズムの最良の特徴を組み合わせ、トランザクションサイズとデータベース内のアイテム数に対して優れたスケールアップ特性を持つハイブリッドアルゴリズム、AprioriHybridを提案した。近年では、RFM分析

を用いた顧客分類とアソシエーション分析を組み合わせる研究が盛んに行われている。亀岡他（2016）は、分析対象を特定の顧客グループに限定することで、通常のアソシエーション分析では見られなかったアソシエーションルールを抽出できるとして、条件付きアソシエーション分析を提案した。これにより、分析対象を層別することで、詳細な顧客理解の可能性を示唆した。しかし、これまでの研究では簡易的な層別による分析に留まり、分析対象の層別に焦点を当てて新たな顧客理解を得た研究は少ない。

そこで本研究では、分析対象の層別に注目し、製品別に顧客を分類して季節性を考慮したアソシエーション分析を行うことで、顧客の購買行動が季節によってどのように変化するかを明らかにする。時系列を考慮したアソシエーション分析を行うことで、時期によって変化する製品への分析が可能となり、各製品別に四半期ごとのアソシエーションルールを抽出する。さらに、アソシエーションルールを基にクラスタ分析を行うことで、異なるクラスタが示す異なる購入パターンやトレンドを特定できる。これにより、優良顧客や特定の購買行動を持つ顧客へのマーケティング戦略を最適化し、季節ごとの購買パターンの違いをより明確に把握することが可能となる。

II. 研究方法

研究方法として、法人を顧客とした印刷機器販売代理店の購買履歴データを用いて研究を実施する。研究対象の期間は2013年5月7日から2023年1月18日であり、63002件、3945社の顧客を対象として分析を行う。本研究は主に3Dプリンタを扱った購買履歴データを使用する。

はじめに、製品種別のRFM分析を行う。次に、RFM分析によって分類したデータから、アソシエーション分析を用いて、優良顧客の四半期毎のアソシエーションルール抽出を行う。その後クラスタ分析を用いて四半期毎にみられる季節性を確認する。

本研究は購買履歴データの中でも主要製品を対象に分析を行い、その主要製品をA社製品と表記する。

3Dプリンタとその関連製品について分析を行うにあたり、本研究で多く見られた製品とそのメーカーについて説明を行う。A社製品の主な製品としてプリントコア、クリーニングフィラメント、PETG、サポートフィラメント、などが挙げられる。プリントコアはプリントヘッドに格納されるカートリッジのことであり、そのプリントコアの目詰まりを解消し、最適な状態に保つものがクリーニングフィラメントである。PETGは優れたプリント適性、強靭性、耐薬品性、耐熱性、耐摩耗性を有している。各商品カテゴリの略称は表1に示す。3DAはA社製品の関連製品であり、3DFはフィラメントを表す。

本研究で対象としたデータには477件の商品カテゴリが

あり、提供されたデータには、その他の製品やそれに付随するサービスなどが含まれる。

本研究ではA社製品を四半期毎に分析を行う。各四半期の定義は以下である。

- ・ 第一四半期 (Q1): 4月から6月
- ・ 第二四半期 (Q2): 7月から9月
- ・ 第三四半期 (Q3): 10月から12月
- ・ 第四四半期 (Q4): 1月から3月

III. 分析

1. RFM分析

RFM分析とは、顧客分析の一種であり、最近の購入日 (Recency)、購買回数 (Frequency)、合計購入金額 (Monetary) の3つの指標で顧客をランク分けする手法である。RFM分析で顧客を分類することで、ターゲットを絞った販売促進を行うことができる。

本研究におけるランク分けについては、各指標に関して5が最高値、1が最低値となっている。また一般的に、Rランクが3、Fランクが3、Mランクが3の顧客をR3F3M3と表記する。分析対象期間中の顧客の購買履歴からR、F、Mをそれぞれ抽出した結果、R5F5M5に該当するA社製品の優良顧客は109名であった。

2. アソシエーション分析

アソシエーション分析とは、顧客が複数の商品を同時に

表—1 各製品カテゴリ

カテゴリ	名称	カテゴリ	名称
3DA-PC	プリントコア	3DF-PVA	ポリビニルアルコール
3DA-CF	クリーニングフィラメント	3DF-SF	サポートフィラメント
3DA-GP	ガラスプレート	3DF-TPU	熱可塑性ポリウレタン
3DA	その他アクセサリ	3DF-NL	ナイロンフィラメント
3DF-PETG	ポリエチレンテレフタレート	3DF-ABS	熱可塑性樹脂
3DF-PLA	ポリ乳酸		

表—2 RFM分析の結果

ランク	範囲	人数
R1	[903,1315)	18
R2	[491,903)	326
R3	[370,491)	170
R4	[248,370)	174
R5	(0,248]	306
F1	[0,8]	226
F2	(8,13]	41
F3	(13,17]	84
F4	(17,25]	55
F5	(25,34]	261
M1	[0,264041]	235
M2	(264041,396062]	42
M3	(396062,528083]	68
M4	(528083,792124]	40
M5	(792124,1056166]	231

購入可能な状態で、どの商品との組み合わせに関連が深いかを数学的に分析する手法である。本研究のデータからは、「Aを購入する→Bを購入する」というルールを抽出できる。ここでのAやBをアイテムと呼ぶ。アソシエーション分析で重要な概念として、支持度（Support）、確信度（Confidence）、リフト値（Lift）がある。支持度は全体顧客の中で、現在注目している商品を購入した顧客の比率を意味する。確信度は商品間の相関を図るときに有効で、商品Aを購入した顧客が、商品Bを購入する割合を示す指標である。リフト値は商品Aと商品Bの因果関係を表し、何らかの因果関係があると1より大きくなる。したがって、リフト値が1より大きければ、「Bを買う」確率が「Aを買うこと」によって上昇していることを示している。

アソシエーション分析の各指標に関する定義は以下である。

- ・ 支持度

$$Support(A \rightarrow B) = P(A \cap B)$$

- ・ 確信度

$$Confidence(A \rightarrow B) = \frac{P(A \cap B)}{P(A)}$$

- ・ リフト値

$$Lift(A \rightarrow B) = \frac{Confidence(A \rightarrow B)}{Expected\ Confidence(A \rightarrow B)} = \frac{P(A \cap B)}{P(A)P(B)}$$

3. クラスタ分析

クラスタ分析とは、データを類似性に基づいて複数のグループに分類する手法である。同じクラスタに属するデータは、類似の特性があり、大量のデータから意味あるパターンや関係性を抽出することができる。特に本研究で用いるk-means法は、Macqueen（1967）が発表し、大規模データに対して効率的であり、指定したクラスタ数に応じてデータを分けるアルゴリズムとして広く利用されている。

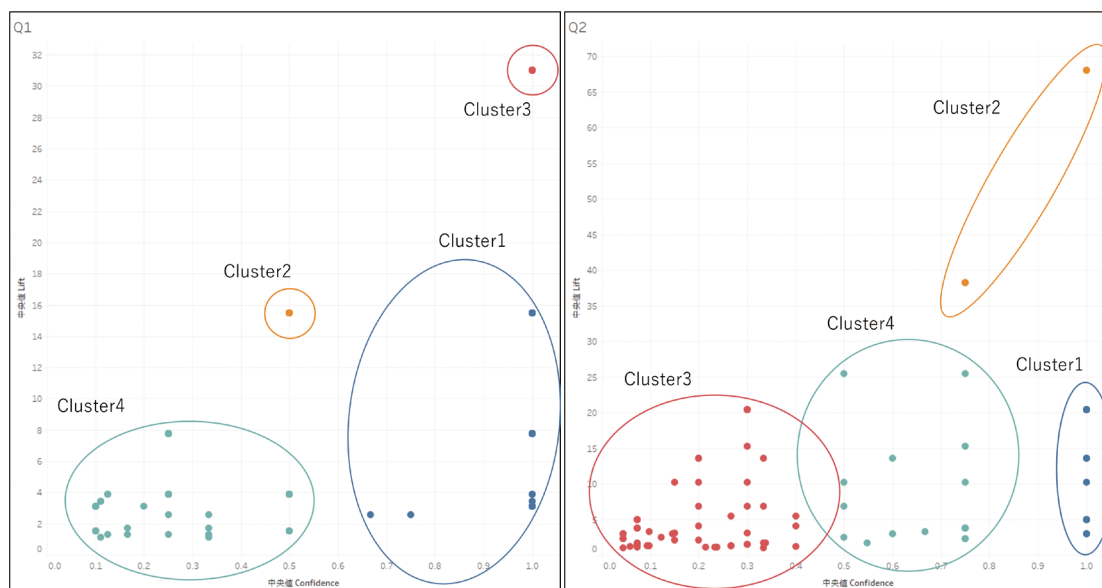
本研究ではアソシエーションルールの確信度とリフト値をプロットすることで、信頼性と有用性の両方を評価し、優良顧客や特定の購買行動を持つ顧客へのマーケティング戦略を最適化する。また確信度とリフト値が高くなるにつれて信頼性が高くなり、他のアソシエーションルールよりも有用性が高いといえる。

各四半期のクラスタ分析の結果を以下に示す。

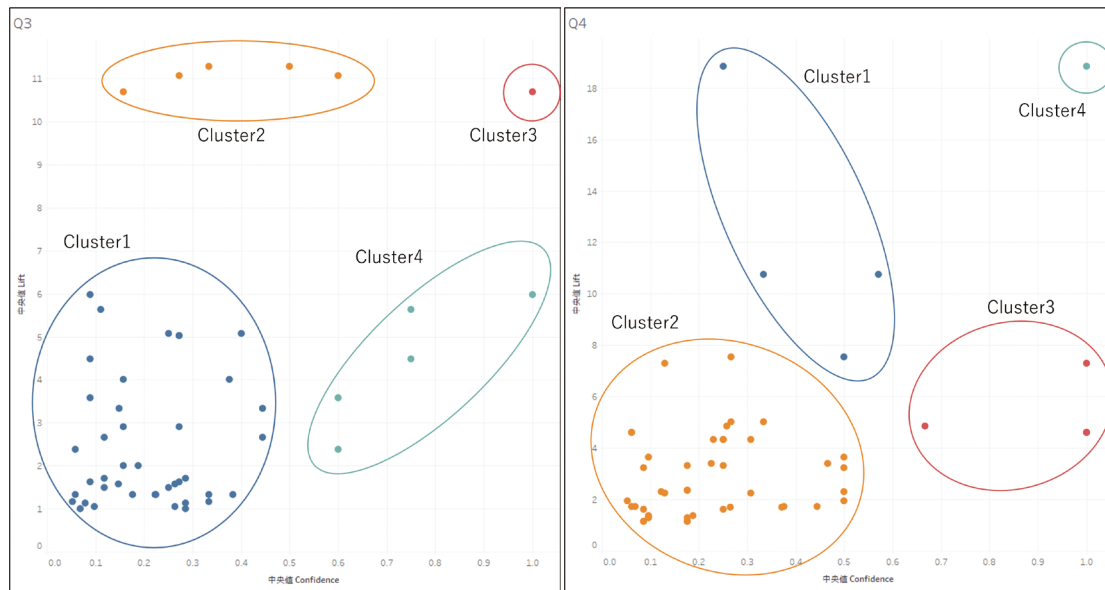
第一四半期（Q1）のクラスタ3では確信度、リフト値ともに非常に高く、強い関連性がみられる。関連製品である3DAを購入する顧客が、ガラスプレート（3DA-GP）などの他の特定の製品も高確率で購入する傾向が強いことを示している。これにより、特定のフィラメント同士の強い依存関係が存在することが示唆された。

第二四半期（Q2）のクラスタ2でも確信度、リフト値ともに非常に高く、強い関連性がみられる。プリントコア（3DA-PC）やサポートフィラメント（3DF-SF）を中心とした複数のフィラメントの組み合わせが多く、一部のルールが非常に強い

図—1 クラスタ分析結果（Q1, Q2）



図—2 クラスタ分析結果 (Q3, Q4)



関連性を持っている。特に夏季において、これらのフィラメントの需要が高いことが示唆された。

第三四半期 (Q3) のクラスタ1では中程度の関連性が見られた。秋から冬にかけて耐熱性やサポート材の需要が高いことが反映されており、特定のフィラメントの組み合わせが強く関連していることが示唆された。

第四四半期 (Q4) のクラスタ4では確信度、リフト値ともに非常に高く、強い関連性がみられた。冬季には、クリーニングフィラメントと高強度PLAの需要が反映されており、多様なフィラメントの組み合わせが強く関連していることが示唆された。

IV. まとめ

本研究では、RFM分析を用いて顧客を分類し、アソシエーション分析を実施した後、クラスタ分析を行うことで各クラスタにおけるアソシエーションルールを抽出することを目的に分析を行った。四半期毎にデータを分析し、特定期間におけるクラスタごとのアソシエーションルールの違いを明らかにすることで、シーズン毎のクラスタ別需要予測を実現した。

本研究の結果を活用することで、購買機会の損失の可能性を低減するだけでなく、優良顧客に向けたクロスセル戦略への活用が期待される。各クラスタの詳細な分析により、特定の時期に特定のフィラメントがどのように関連付けられて購入されるかを把握することができるため、効果的なマーケティング戦略の構築が可能となる。

今後の課題として、予算消化のタイミングと季節性を判別することが難しいため、業種等の顧客属性別に詳細な分析を行う必要がある。また、各クラスタの特徴をさらに深掘りし、より具体的な購買パターンを明らかにすることで、さらに精度の高い需要予測とマーケティング戦略の提案が求められる。これにより、顧客満足度の向上と売上の最大化が期待できる。

謝辞

本研究の遂行にあたり、指導教員として終始多大なご指導を賜りました、千葉工業大学大学院 社会システム科学研究科 マネジメント工学専攻 加藤和彦教授、武田善行准教授に深く感謝いたしますとともに御礼申し上げます。

引用文献

- 五十嵐千人・松井佑介・南弘征・水田正弘 (2015). 「共同購入型クーポンサイトにおける顧客の分類及び群ごとの購買傾向の解析」『計算機統計学』28(2), 139-146.
- Agrawal, Rakesh, Srikant, Ramakrishnan (1994). 「Fast algorithms for mining association Rules」『In Proc. of the 20th int. conf. on very large data bases』1215, 487-499.
- 亀岡瑠・船山貴光・宗像昌平・山田実俊・八木圭太・山本義郎 (2016). 「条件付きアソシエーションルールによる顧客の購入特徴の抽出」『計算機統計学』29(1), 57-64.
- Macqueen, James (1967). 「Some Methods for Classification and Analysis of Multivariate Observations」『Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability』1, 281-297.

阿部誠 (2014). 「RFM データを用いた顧客生涯価値の算出」『マーケティングジャーナル』34(1), 73-90.

山田実俊・山本義郎 (2016). 「対応分析を用いたアソシエーションルールによるアンケート結果の可視化」『日本分類学会』5(1), 3-15.