

6N-7

時系列データから自動発見されるパターンにもとづく 推論のPOSデータ活用への応用

阿部哲也[†]* 永島繁子[†] 前田茂[†] 長野重美[†]

(株) 東芝 [†]マルチメディア技術研究所 [‡]流通・金融・情報システム事業部

1. はじめに

学習用の大量データから、そこに存在するパターンを自動的に発見・学習しておき、対象とするデータがどのパターンに近いかを調べることから指示や警告、予測を推論するパターン認識型推論の方式を検討した。そして、実際のスーパーマーケットのPOSデータ(Point of Sales) 活用に応用して、「ロス」の傾向を指摘し、発注量の増減を指示するシステムを試作したので報告する。

2. 自動発見されるパターンによる推論

2. 1. パターン認識型推論の概要

検討したパターン認識型推論方式の概要を図1に示す。学習用データからパターンを発見・学習しておき、対象とするデータについて、「あのパターンだから、こう警告しよう」と推論することが目的である。[1]

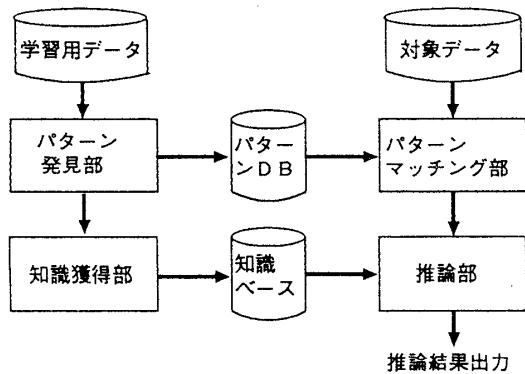


図1 パターン認識型推論の基本構成

学習時は、学習用データに内在するパターンを自動的に発見し、参照パターンとしてパターンDBに格納する。そして発見したパターンに応じた知識を獲得して知識ベースに格納する。

*An application of the automatically discovered pattern based reasoning to the practical use of P.O.S. data, Tetsuya ABE[†], Shigeko NAGASHIMA[†], Shigeru MAEDA[†], Shigemi NAGANO[†], [†]Multimedia engineering lab., [‡]Distributing, banking & information systems div., Toshiba corp., 70, Yanagi-cho, Saiwai-ku, Kawasaki 210, Japan

推論時は、対象データをパターンDB中の参照パターンとマッチングし、どの参照パターンに似ているかを調べる。そして、推論部が、その意味するところを知識ベース中の知識で推論し、警告や注意などを出力する。

2. 2. 構成各部の役割と仕組

(1) パターン発見部

学習用データをクラスタリングし、発見された各クラスターをパターンと捉えてその情報をパターンDBに格納する。図2は、5つの商品の日々の売上個数の推移と、発見したパターンを示している。5つの商品の60日間のデータを、60次元のデータ5サンプルと捉えてクラスタリングし、各クラスターの平均値を代表パターンとして表示している。

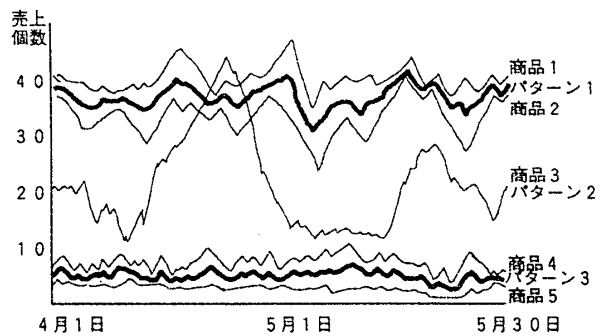


図2 パターン発見の例

(2) 知識獲得部

パターン発見部が発見・表示したパターンに応じて、「このパターンに対しては、こういうことが言える」といった知識の記述を行ない、知識ベースに格納する。

(3) パターンマッチング部

推論対象データを参照用パターンとパターンマッチングし、その結果を確信度付きで推論部に送る。

(4) 推論部

パターンマッチング結果と知識から、対象データについての注意や警告、指示、予測などを推論する。知識をルール形式で記述しておき、パターンマッチング部から

確信度付きで結果が入力されれば、ルール推論によって確信度付きの推論結果が得られる。

3. POSデータ活用への応用

3.1. ロス傾向指摘システム

パターンにもとづく推論の考えをスーパーマーケットのPOSデータ活用に応用して「ロス傾向指摘システム」を試作した。スーパーマーケットにおける破損破棄や機会損失といった「ロス」の傾向を指摘し、注意やアドバイスを出した上で、商品の発注量の増減指示を出すシステムである。

スーパーマーケットでは、商品が売れ残って破棄されたり、売り切れて商売の機会を損失するといった「ロス」を少なくすることが重要である。そこで、POSデータを活用し、日々の売上や仕入れ及び「ロス」の傾向を自動的に指摘して、適切な発注量の指示を推論できるシステムが望まれている。

3.2. システムの試作

図1の基本構成を商品発注のドメインに対応させて「ロス傾向指摘システム」を試作した。データとしては、仕入れと売上の差を取った売れ残りの日々の推移を1つの時系列サンプルとした。

(1) パターン発見部

売れ残りの推移について、クラスタリングによって、いくつかの代表パターンを発見して表示するとともに、パターンデータとして蓄える。

(2) 知識獲得部

パターン発見部で発見・表示された代表パターンを見て、各パターンに応じた知識(ルール形式)を人手で記述した。ルールの例を以下に示す。図中、数値は確信度を示す。

```
(defrule ルール6
  (same 対象パターン パターン6)
  (same 商品分類 米飯弁当)
  ==>
  (発注指示 減少 0.4)
  (機会損失 あり -0.6)
  (パターン傾向 "いつも少し売れ残る傾向" 0.8)
  (アドバイス "少し発注を減らしましょう" 0.8)
)
```

図3 ルール例

(3) パターンマッチング部

対象商品の売れ残りデータの推移と、学習してある代表パターンとをパターンマッチングする。距離が近いほど類似性が高く確信の度も高いとみなし、距離から変換した確信度をつけて結果を出す。

(4) 推論部

ルール型知識とパターンマッチング結果を用いて、指摘やアドバイスを後ろ向き推論して利用者に提示する。また、パターンマッチング部からは対象データがどのパターンに近いかわかるので、ルール自体の確信度と合わせて確信度計算し、発注の増減指示を確信度付きで提示する。

3.3. システム動作

実験として、約50個の商品の実データを用い、120日分のデータから代表パターンを発見させ、その後の30日分のデータに対して推論を行なった。その結果、いつも売れ残るものや売り切れが多いものなど、数種類の代表パターンを数秒程度で発見できた。そして、約50個の商品について10秒程度でアドバイスや発注増減指示を推論・表示できるようになった。図4に推論結果表示画面の例を示す。

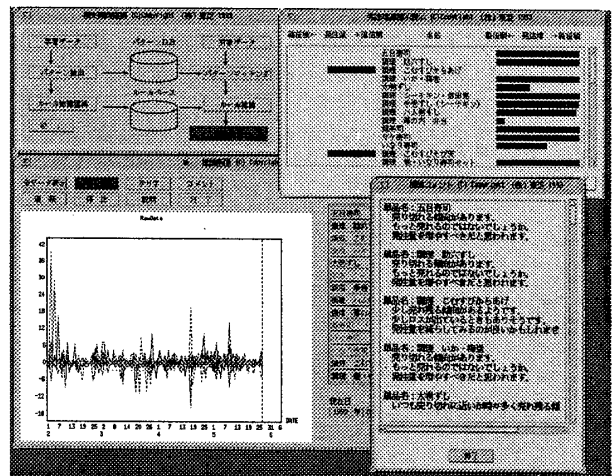


図4 推論結果表示画面

4. おわりに

以上、自動的に発見されるパターンにもとづく推論の考えを示し、それをPOSデータに適用した「ロス傾向指摘システム」について報告した。パターン発見手法や知識獲得支援などの機能向上を目指すとともに、試作システムの実用レベルへの向上や他分野への応用も検討したい。

参考文献

- [1] J.R.Whitely, J.F.Davis, "Qualitative Interpretation of Sensor Patterns," IEEE EXPERT, Apr., 1993