

4N-1

知識ベースを用いた株式の業種分類

藪内 康宏

株式会社パナウエル システム部

1. はじめに

株式投資において、ポートフォリオの構築に個別銘柄がどの業種のグループに所属しているかという情報が非常に重要視されている。それは次のような観点に基づくものである。まず第1に、リスク分散投資という視点に立った場合、同じ様な株価変動を示す銘柄に投資するよりも異なった株価変動を示す銘柄に投資する方がリスク分散効果が高くなる。それゆえ、株価変動に基づく業種のグループを構成することにより、リスク分散効果を狙った銘柄選択が容易になる。株価変動の類似性に基づいた業種分類の例には、クラスター分析を用いた業種分類[1]が挙げられる。第2に、企業の業態についての情報は個別銘柄を選別する上で重要な判断材料を提供してくれる。この業態に基づいた業種分類としては東証28業種分類が利用されている。しかしながら、この分類では業種が28と固定されているため、様々な投資家の要求に対して柔軟な対応ができないという問題点がある。本稿では、第2の視点に立ち、投資家の幅広い要求に対応できるような、企業業態に基づいた業種分類システムを、知識ベースを利用することによって実現する。

2. 業種分類システム

業種分類システムは、分類対象となる銘柄のデータ等から作成された知識ベースと、システムのメイン処理部である業種分類操作から構成されている。図1にその概略を示し、以下に知識ベースと業種分類操作について説明する。

2.1 知識ベース

知識ベースは、売上構成品目モデル、株式モデル、及び株式銘柄インスタンスから構成されており、それぞれMinskyのフレーム[3]を用いて表現されている。売上構成品目モデルは、売上構成品目(以下では品目)の類似性を表現するために作られたモデルである。そして、品目の類似関係をネットワークで表現している。図2にモデルの例を示す。ここで類似性を表す指標として類似度を導入する。この類似度は枝の距離によって定められ、距離が近いほど類似性が高くなり類似度は小さく、

逆に距離が遠いほど類似性が低くなり類似度は大きくなる。株式モデルは、全株式銘柄にとって共通の知識や推論が記述されているモデルである。株式銘柄インスタンスは、株式モデルのインスタンスの集まりである。それぞれのインスタンスには、日経会社情報[4]の銘柄コード、銘柄名、売上構成という属性情報が格納されている。

2.2 業種分類操作

株式銘柄の業種分類を以下の2段階の操作で行う。まず第1段階として、全銘柄について各々の類似銘柄を集めた類似銘柄群を推論する。各銘柄に対する類似銘柄の選択は、銘柄の売上構成品目と売上構成比率に注目し、これに適切な類似度を導入することにより、売上構成品目と売上構成比率のパターンマッチングによってなされる。

第2段階では、マッチングによって得られた類似銘柄群において、互いに同一業種である銘柄を確定する。ここで任意の2銘柄(A, B)が、同一業種である為の定義を導入する。BはAの類似株式であるということを記号 $A > B$ で書き、次のように定義する。

$$A > B \Leftrightarrow B \in [A]$$

ここに、[A]とはAから生成される類似銘柄群を表す。更に、関係>に関して以下のような推論が成立すると仮定する。

$$A > B \wedge B > C \Rightarrow A > C$$

このとき、AとBが同一業種のグループに所属可能であることを記号 $A \sim B$ で書き、次のように定義する。

$$A \sim B \Leftrightarrow A > B \wedge B > A$$

この定義から関係 $\sim$ は以下の3条件を満足し、同値関係となる。従って、関係 $\sim$ に基づくグループ分けはうまく定義される。

[条件1] $A \sim A$

[条件2] $A \sim B \Rightarrow B \sim A$

[条件3] $A \sim B \wedge B \sim C \Rightarrow A \sim C$

A, B, C: 任意の銘柄

この関係 $\sim$ に基づいて、同一業種を抽出することによって業種分類が行われる。

3. システムの応用例

業種分類システムを用いて、1991年12月2日の日経会社情報のデータから日経平均採用銘柄の食品に所属する17銘柄（日本製粉、日清製粉、日本甜菜製糖、森永製菓、明治製菓、明治乳業、サッポロビール、アサヒビール、麒麟ビール、宝酒造、合同酒類、メルシャン、ホーネンコーポレーション、日清製油、キッコーマン、味の素、ニチレイ）を対象に業種分類を行った。分類は、類似性の高い類似度1から類似性の低い類似度5に対して行った。表1に結果を示す。類似度1の時は13組、2の時は8組、3の時は5組、4の時は2組、5の時は1組という結果を得た。この表より、類似度が小さいほど分類結果は細かくなり、類似度が大きいほど分類結果は粗くなることわかる。特に類似度1と2では、類似銘柄群の推論の際に売上構成目目のマッチングの成立しない銘柄が多かったために、グループの要素が単一銘柄となるものが多かった。逆に、類似度5では売上構成目目のマッチングが全て成立したために、グループは一つとなった。これは、売上構成目目モデルの最大類似度が5となっていたからである。

4. おわりに

本研究では、知識ベースを用いた新しい株式の業種分類システムを提供した。本システムを用いれば、分類条件の類似度を変化させることによって、グループ数を容易に調節することが可能である。また、売上構成目目モデルをより“優秀”な投資家が作成することによって、より理想的な分類結果が得られるであろう。今回は、日経平均採用銘柄の食品に所属する17銘柄についての業種分類を行ったが、今後は、全ての個別銘柄を本システムを用いて業種分類し、その結果を評価するための評価基準を設定し、既存の業種分類方法との比較を行い、有効性を検証する。

5. 参考文献

- [1] 菅原周一、片岡 淳、株式ポートフォリオ運用のための業種分類、証券アナリストジャーナル 1992. 7, p p. 84-93.
- [2] Farrell, James L. Jr., Homogeneous Stock Groupings, Financial Analysts Journal, May/June 1975, pp.50-62.
- [3] Winston, Patrick H. & Horn, Berthold K. P., 白井良明、安部憲広訳, L I S P, 培風館
- [4] 日本経済新聞社, 日経会社情報, 季刊1992-1新春号

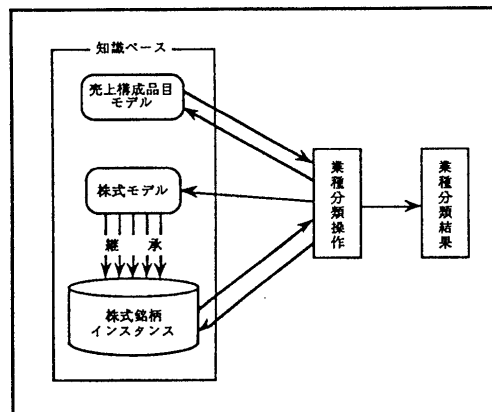


図1. 業種分類システム

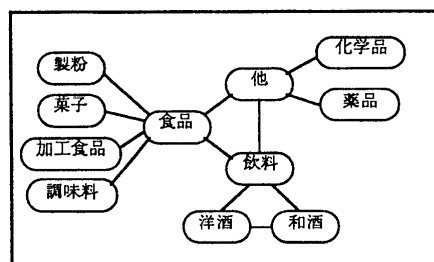


図2. 売上構成目目モデル

表1. 分類結果

類似度	分類内容	総グループ数
1	(日本製粉、日清製粉) (日本甜菜製糖) (ニチレイ) (日清製油、味の素) (明治製菓) (ホーネンコーポレーション) (キッコーマン) (森永製菓) (明治乳業) (サッポロビール、アサヒビール、麒麟ビール) (メルシャン) (合同酒類) (宝酒造)	13
2	(日本製粉、日清製粉、日本甜菜製糖) (ニチレイ) (日清製油、味の素、明治製菓) (ホーネンコーポレーション、キッコーマン) (森永製菓) (明治乳業) (サッポロビール、アサヒビール、麒麟ビール、メルシャン) (合同酒類、宝酒造)	8
3	(日本製粉、日清製粉、日本甜菜製糖、ニチレイ) (日清製油、味の素、明治製菓) (ホーネンコーポレーション、キッコーマン) (森永製菓、明治乳業) (サッポロビール、アサヒビール、麒麟ビール、メルシャン、合同酒類、宝酒造)	5
4	(日本製粉、日清製粉、日本甜菜製糖、ニチレイ、日清製油、味の素、明治製菓、ホーネンコーポレーション、キッコーマン、森永製菓、明治乳業) (サッポロビール、アサヒビール、麒麟ビール、メルシャン、合同酒類、宝酒造)	2
5	(日本製粉、日清製粉、日本甜菜製糖、ニチレイ、日清製油、味の素、明治製菓、ホーネンコーポレーション、キッコーマン、森永製菓、明治乳業、サッポロビール、アサヒビール、麒麟ビール、メルシャン、合同酒類、宝酒造)	1