

4N-5

ニューロと統計手法を併用した
債券格付推定モデルの提案

牧本 伸生・ 増井 裕也・ 神成 広之** 吉原 郁夫*

* (株)日立製作所 システム開発研究所

** (株)日立マイコンシステムズ

1. はじめに

債券格付(AAA, AA+, ..., など)は企業の債券の優良性を表す指標であり、第三者機関により決定される。経営方針等の定性的な情報も格付に反映されるが、主な格付決定要因は財務指標である。本報では財務指標のみから格付値の推定を行なうモデルを提案する。

推定には多層型ニューラルネットワークを用いる。提案モデルの特徴として、第1に多変量解析を用いてニューロへの入力データを前処理する。従来このような前処理はあまり重視されていなかったが、大量異種データからなる実問題に対してはこれにより本質的データを抽出して学習の容易化を図ることが必須である。

第2に格付及び業種を勘案し、銘柄を幾つかのグループに分割して、複数のネットワークを使用する。これによって、単一のネットワークのみを使用する場合に比べて学習の容易化が期待できる。またクラスタリングにより類似データを積極的に統合することで、学習サンプル不足による汎化能力低下が防止できる。

2. モデル構成

提案モデルは、以下に示す3つの処理部から構成される(図1)。

- (a) 前処理部: まず入力された財務指標(営業利益など数種)を、格付との非線形回帰分析・財務指標同士の主成分分析に基づいて推定に有効な特徴量に変換する。次にそれを正規化することにより、ニューロへの入力に適した形式にする。なお、ここで用いるパラメタ(主成分行列等)は予め求めておく。
- (b) セレクタ部: 「(c) 推定部」は格付及び業種に応じた8個のネットワークからなるが、その中のどれを推定に用いるかをここで決定する。そこでまず、「(a) 前処理部」の出力から重回帰分析(財務指標と格付値)に基づき格付をラフに推定する。この結果と予め決められた業種分類とを用いて、ネットワークの一つを選定する。業種分類とは、AIC(Akaike's Information Criterion)を利用して銘柄群を予めクラスタリング(小業種を小数の大業種に統合)しておいたものである。
- (c) 推定部: 「(a) 前処理」の出力を「(b) セレクタ」で決定したニューラルネットワーク(3層型)に通し、その出力パターンから最終的な格付推定値を決定する。なお、各ネットワークの中間層ユニット数は、AICの値が最小になるように予め決めておく[1]。また結合係数も、誤差逆伝播法により実データを学習させ、決定しておく。

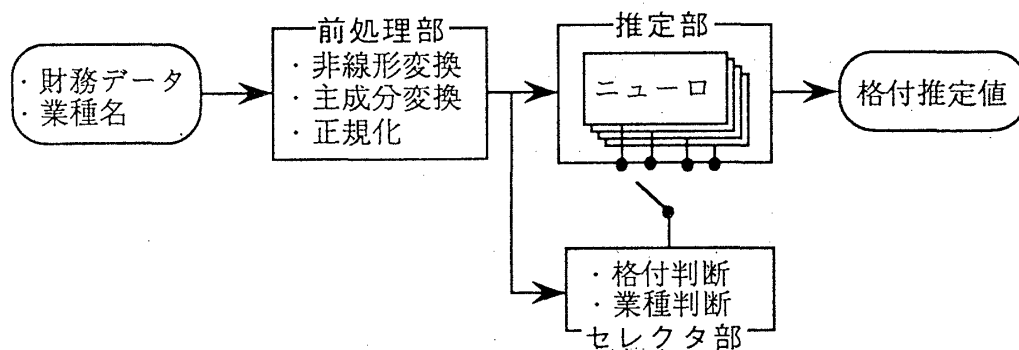


図1 システムの全体構成

A proposal of a bond rating system by neural network and statistical model

Nobuo MAKIMOTO*, Hironari MASUI*, Hiroyuki KAN-NARI**, Ikuo YOSHIHARA*

*Systems Development Laboratory, Hitachi, Ltd. **Hitachi Micro Computer Systems, Ltd.

3. AICを利用したクラスタリング方式

学習の容易化並びに汎化能力の向上を図るために、業種のクラスタリングを行なう。このため本報では、AICを利用したクラスタリング方式を提案する。セレクト部の業種判断ではこの結果を用いている。

AICは統計モデル（ここでは小業種を大業種へ統合する組合せ；業種群と呼ぶ）の対数尤度（その-2倍のある種の不偏推定量）である[2]。従って、各ネットワークの学習誤差から求められるAICの総和（以下総AICと呼ぶ）が小さいほど、全体として良い業種群であるといえる。

クラスタリングの手順を以下に示す。総AICが最小になる業種群を、Simulated Annealingと等価な確率的手続きに従い探索する（図2）。

step 1 初期設定／更新

- ・SAの温度パラメタ更新
- ・業種群仮更新（ランダムに1小業種を選んで、その属する大業種を仮に変更）

step 2 学習

- ・仮大業種毎に各ネットワークで学習

step 3 評価・更新

- ・各ネットワークでAICを求め、総AICを評価
- ・仮更新前の業種群の総AICの値と比較し、その差及び温度に応じた確率で業種群を実際に更新
- ・終了判定（終了条件不成立の場合 step 1 へ）

4. 実験条件及び結果

提案モデル及びそのサブモデルを用いて格付推定シミュレーションを行なった。各モデルの正答率を表1に示す。学習用にはある年度の約400社分のデータ（財務指標+格付）を用いた。

正答率評価は、学習に用いたデータ及び別年度のデータに対して行なった。モデル1では、統計解析で格付の大まかな範囲を推定でき、これに基づくネットワーク選択が可能であることを確認した。モデル3（提案）では、モデル2と比較していずれのデータに対しても高い正答率が得られた。これは、複数小ネットワークの使用により学習能力・汎化能力ともに向上したことを示している。

5. おわりに

実験結果から、提案モデル（統計解析による本質的情報の抽出、レンジ分割・クラスタリングによる問題の分割）が格付推定問題に有効であることが分かった。

6. 謝辞

本研究にあたりご指導頂いた千葉大学の松葉育雄教授に感謝致します。

参考文献

- [1] Matsuba: IJCNN Singapore, pp. 2140-2143 (1991).
- [2] 坂元ほか：『情報量統計学』、共立出版(1983)。

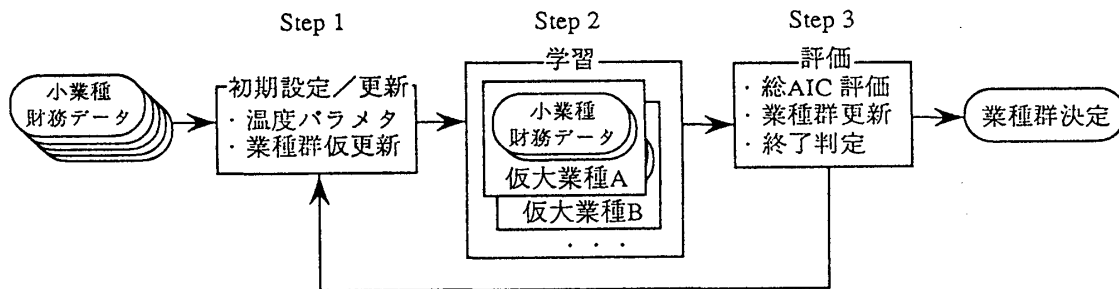


図2 業種群分類のクラスタリング手順

表1 格付推定シミュレーション結果

No.	モデル	学習年度データ	別年度データ
1	統計解析（非線形回帰・主成分分析・重回帰）	46%	45%
2	統計+ニューロ（単一）	51	37
3	統計+ニューロ（レンジ分割・クラスタリング）	81	60