

平成 30 年 5 月 21 日現在

機関番号：32660

研究種目：若手研究(B)

研究期間：2015～2017

課題番号：15K21395

研究課題名(和文) 倒産・上場廃止予測にとって最適な財務分析指標のパターン認識的選出

研究課題名(英文) Pattern-recognition-based selection of optimal financial indicators for bankruptcy / delisting prediction

研究代表者

保坂 忠明 (Hosaka, Tadaaki)

東京理科大学・経営学部経営学科・講師

研究者番号：60516235

交付決定額(研究期間全体)：(直接経費) 1,500,000円

研究成果の概要(和文)：本研究では、企業の倒産予測に関して、1) 予測に有効な財務指標の抽出、および倒産企業と継続企業の識別に役立つ判定式の導出を一貫した枠組みの中で実現する方法の提案、2) 様々な分野で話題となっている深層学習(ディープラーニング)を用いて予測を行う方法の提案を行った。結果として、1) 倒産の1年以上前でも高い精度で予測が可能である、2) 深層学習を利用した手法は従来手法と比較して高い精度で予測が可能である、ことが示された。

研究成果の概要(英文)：In this research, we aim to resolve some problems with respect to corporate bankruptcy prediction. We propose methods of 1) realizing the extraction of financial indicators and the derivation of discriminant functions in a consistent framework, and 2) applying the techniques of deep learning to bankruptcy prediction. As a result, we have shown that it is possible to predict with high accuracy even more than one year before bankruptcy. It was also shown that the method using deep learning can predict with high precision compared with the conventional methods.

研究分野：機械学習を利用した財務分析

キーワード：倒産予測 実質破綻予測 機械学習 AdaBoost 深層学習 画像化 畳み込みニューラルネットワーク

### 1. 研究開始当初の背景

機関、個人を問わず投資家が証券取引において利益を生み出すためには、現在の経営状態から企業の今後の業績を可能な限り正確に予測・予知することが重要である。その代表的な例の一つは、企業の倒産（実質破綻）やそれに伴う上場廃止の予知である。

企業の倒産予知に関しては、本研究の開始以前にもパターン認識や機械学習を用いて多くの研究が報告されてきた。倒産予知に関する多くの従来研究では、会計的な視点やそれまでの研究の実績に基づいて数個から5個程度の財務指標があらかじめ設定され、予知モデルが構築されている。そのため、予測・予知にとって潜在的に適した財務指標がそもそも分析対象から除外されている可能性があった。また、財務指標の選択と機械学習等による予知モデルの構築が別々のプロセスとなるため、手法全体としての予知精度の最適性が保証されない問題もあった。

また、本研究を開始する頃には、深層学習（ディープラーニング）が、機械学習や人工知能の分野だけでなく一般社会においても大きな注目を集めていた。しかし、深層学習の工学分野への応用の多様性と比較すると、財務分析への応用は株価の上昇・下降の予測に適用されたいくつかの例が知られている程度であり、非常に限定的であった。

### 2. 研究の目的

前節で述べた企業の倒産予知に関する問題点の解決を中心として本研究の目的を設定した。具体的には、(1)倒産予知における財務指標の選択と予知モデルの構築を単一の最適化の枠組みの中で扱う、(2)深層学習を利用して倒産予知を実現する方法を提案する、ことを目的とした。ただし、目的(1)については、早期の段階での倒産予知精度の向上を念頭においており、目的(2)については手法の提案にとどまらず、従来手法を上回る予知精度を実現することを目指すこととした。

以上の目的(1)を達成するためにAdaBoostと呼ばれる機械学習手法を利用した。また、目的(2)を達成するために複数の財務指標をまとめて一枚の画像として表現して深層学習の入力として利用することを考えた。

### 3. 研究の方法

本研究の基本的な問題設定は、財務指標を頼りに倒産企業（実質破綻企業）と継続企業（存続企業）の識別を行うことである。機械学習によってこのような二クラス分類問題を実行するためには、対象となる二クラスに該当する企業のデータが必要となる。本研究では、倒産企業として、2002年1月から2016年6月までに日本の株式市場である東証、大証、旧ナスダック・ジャパンスタンダード、旧ヘラクレススタンダード、旧ヘラクレスグ

ローズ、旧ジャスダックのいずれかの市場で経営破綻（もしくは実質的にそれとみなせる事由）によって上場を廃止した100社程度の企業を対象とした。なお、企業数は欠損値処理の仕方などに影響を受けて変動するため、研究の初年度から最終年度にかけて若干の差異が生じている。なお、実質破綻とみなすことができる上場廃止の理由としては、1)債務超過、2)銀行取引の停止、3)破綻・再生・更生手続き、4)事業活動の停止（ただし合併を除く）とした。一方で、本研究で対象とする継続企業は、2016年6月時点で東京証券取引所一部、または二部に上場していた企業とした（同様に多少の変動があり、2062~2287社である）。

これらの企業の連結決算における4期分の貸借対照表、損益計算書の二種類の財務諸表を日経NEEDS Financial QUESTより取得した。貸借対照表には175項目、損益計算書には88項目が含まれている。

以下に、研究目的の(1)、(2)を実現するための方法について、それぞれに分けて記す。

#### (1) 単一の最適化の枠組みに基づく財務指標選択と予知モデルの構築

本研究では大きく分けて二種類の財務指標を生成した。ひとつは、単一の会計年度の財務諸表から生成される財務比率である。同一の会計年度の貸借対照表、損益計算書の中から任意の二項目を選び（二つの財務諸表から一項目ずつを選択する場合も含む）、それらの比率を求めた。もう一方の財務指標は、二つの会計年度の財務諸表を組み合わせ生成される財務比率である。比率の生成方法には複数のパターンが存在するが、それらはおもに財務項目の年度間の変化を反映するものである。倒産時点から遡った決算データであっても複数年度の財務諸表から生成される財務比率を使うことで、識別精度を向上させて早期予知を実現することを目指している。

財務指標の選択と識別関数の構築には、機械学習の一手法であるAdaBoostを用いた。AdaBoostでは、まず二クラスの識別を行うために最適なひとつの財務比率が抽出される。そして、選ばれた財務比率が誤識別する学習データの重みを増加させる。次のステップでは、新たな重みづけの下で最適な財務比率が抽出され、同様に学習データの重みも更新される。その後は、同様のステップを繰り返すことになる。このアルゴリズムにより、最初の方のステップで識別を誤る「難しい」データに特化した財務比率が後続のステップでは選択されやすくなる。

さらに、AdaBoostでは選択された財務比率群の重み付き多数決によって最終的な識別関数を表現できることが、ある種の最適化問題の解として理論的に導出される。これより、

AdaBoost は、財務指標の選択と識別関数の構築を単一の最適化の枠組みの中で扱っているとされる。

継続企業に関しては全 2287 社の中から 94 社を無作為に選択し、継続企業 94 社と倒産企業 94 社（この研究を行っていた時点での企業数）を用いて評価実験を行った。この継続企業 94 社と倒産企業 94 社を 1 セットとして扱い、継続企業の組み合わせを 100 回変えて、それら 100 セットの各々に対して評価をした。

各セットの評価実験においては、1)188 社の企業データのうち、187 社分を学習データセットとして、AdaBoost により財務指標選択と識別関数の構築を行う、2)残りの 1 社のデータを評価データ、すなわちクラスラベルが未知の新規データと想定して倒産企業、もしくは継続企業の識別を行う、といった手順をとった。学習データセットと評価データの組み合わせを変え、全 188 通りのパターンをすべて試した。最終的には全 188 通り×100 セットの試行に渡る平均識別率により提案手法の性能評価を行った。

#### (2) 深層学習を利用した倒産予知

深層学習を利用して大きな成功を収めた研究のひとつとして、畳み込みニューラルネットワークを利用した画像認識問題がある。そこで、本研究では畳み込みニューラルネットワークの財務分析への応用を考えた。ただし、畳み込みニューラルネットワークは、画像に適した手法と言えるため、財務データのような数値データにはそのままでは適用しづらい。

そこで、各企業の各年度の財務データから可能な限り多くの財務比率を生成し、それらを一枚のグレイスケール画像で表現した。そのために、各財務比率に特定の画素位置（画像座標）を対応させ、財務比率の値に応じてその画素の輝度値を決定した。ただし、この対応関係の決定においては、相関の高い財務比率同士がなるべく近傍に配置されるようにした。また、財務比率の値が大きいほど、対応する画素は白に近い色で表現され、逆に財務比率の値が小さいほど、対応画素は黒に近い色で表現されるようにした。

一般に深層学習には多くの学習データが必要と考えられている。本研究では、特に倒産企業クラスのデータが 408 個（102 社 4 期分；この研究を行っていた時点での企業数）しか存在せずに非常に少ないと言える。そこで、倒産企業、継続企業ともに任意の 2 期を選び、すべての財務項目に対して、該当 2 期の数値の加重平均を求めることでデータ数を増大させた。画像認識問題においても、左右・上下の反転、並進操作をはじめとする幾何学的変換等によりデータ数（画像枚数）を

増大させることが頻繁に行われるが、これはそれに準じる方法と言える。

財務データを上記の方法により画像へと変換し、最終的に 5 組の学習データセットと評価データセットを作成した。学習データセットには、倒産企業および継続企業のクラスにそれぞれ 7520 データが属しており、一方で、評価データセットには、倒産企業として 88 データと継続企業として 7928 データが含まれている。これらの学習データセットを用いて畳み込みニューラルネットワークのパラメータを学習させた。本研究におけるネットワークは、Google 社が提案し、2014 年の ILSVRC (ImageNet Large Scale Visual Recognition Competition) で 1 位となったネットワークに基づいている。学習されたネットワークを使い、評価データセットに対する識別率を求めた。5 組の学習データセットと評価データセットを使って実験を繰り返し行い、全 5 回の試行に対する平均識別率により提案手法の有効性を検証した。

#### 4. 研究成果

本節でも二つの研究目的に分けて成果を記す。

##### (1) 単一の最適化の枠組みに基づく財務指標選択と予知モデルの構築

単一の会計期間（倒産の直前期から 3 期前まで）の財務データから生成される財務比率のみを利用した場合の識別率を図 1 に示す。直前期から時間が遡るほど、倒産企業と継続企業を識別することが難しくなることが分かる。また、二つの会計期間の財務データから生成される財務比率を利用した場合の識別率を図 2 に示す。多くの場合に識別関数に用いる財務比率の数を増やしていくと（AdaBoost のステップ数を増やすことに相当）、図 1 に示す単一の会計期間のみの財務比率を用いる場合よりも識別率が向上している。特に、1 期前と 2 期前の財務データから生成される財務比率によって識別関数を構築した場合には、直前期の財務データのみを用いる場合と同程度の識別率が得られた。これは、倒産の 1 年前でも高い精度で予知が可能であることを示している。

1 期前と 2 期前の財務データを組み合わせた場合において、AdaBoost により、最も高い頻度で抽出される三つの財務比率の組み合わせは、

- ・利益剰余金(1 期前) ÷ 流動資産(1 期前)
- ・評価換算差額等(1 期前と 2 期前の差分) ÷ 資本金(1 期前)
- ・特別損失(1 期前と 2 期前の平均) ÷ 資産合計(1 期前と 2 期前の平均)

であった。

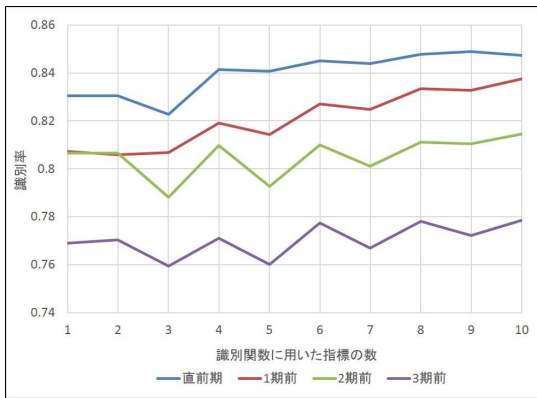


図1：単一会計年度の財務データから生成される財務比率を用いた場合の識別率

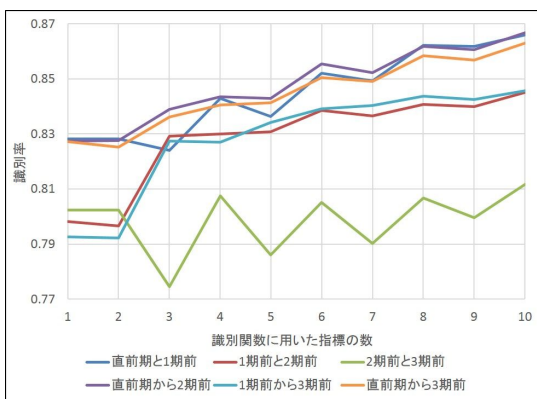


図2：二つの会計年度の財務データから生成される財務比率を用いた場合の識別率

「1期前から3期前」などと3期以上の表現となっている場合には、それらに含まれる任意の2期から生成される財務比率の和集合が対象である。

### (2) 深層学習を利用した倒産予測

ネットワークの層数を変化させた場合の提案手法の性能を図3に示す。なお、ここでは、評価データに含まれる倒産企業と継続企業のデータ数が異なるため、識別率ではなくF値によって性能を示している。また、財務比率の画像化の際に、相関の高い財務比率同士を近傍画素に配置する場合(Correlatedと表示している)と比較するために、財務比率をランダムな画素位置に対応させた場合の性能(Randomと表示)も示している。図3より層数を増やすほど、識別性能が上昇することが分かる。また、相関のある財務比率同士を近傍に配置することは識別にとって有効に働くことが示された。

他の代表的な機械学習を使用した倒産予測、実質破綻予測の手法との比較を行った。ここでは、1)決定木(CART アルゴリズム)、2)線形判別分析(LDA)、3)サポートベクターマシン(SVM)、4)AdaBoost、を用いた方法(本研究の手法(1))を比較手法として取り上

げた。これらの方法によって得られる平均識別率を倒産企業と継続企業に分けて図4、図5にそれぞれ示す。いずれの手法についても、データセットは提案手法で用いたものと同じとした。線色が手法を表している。横軸は、識別関数に取り入れる財務比率の個数を表し、縦軸は5つの評価データセットに対する平均識別率を表している。なお、CARTおよび提案手法については変数選択の処理が含まれないため、横軸に関係なく一定値が示されている。これより、提案手法がいずれの手法と比較しても総じて顕著に優位であることが示された。

提案手法(2)の識別性能は高いものの、提案手法(1)やいくつかの従来手法と異なり、どの財務比率が識別に強い影響を与えているかを提案手法(2)から知ることは困難である。そのため、倒産に至るメカニズムなどを調べるといった目的には提案手法(2)は適さない点を認識しなければならない。

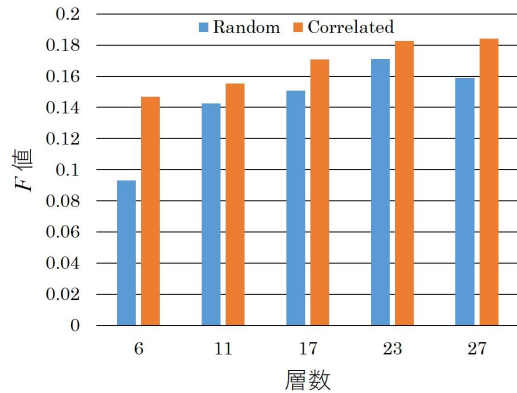


図3：ネットワークの層数に対する識別性能(F値)の変化

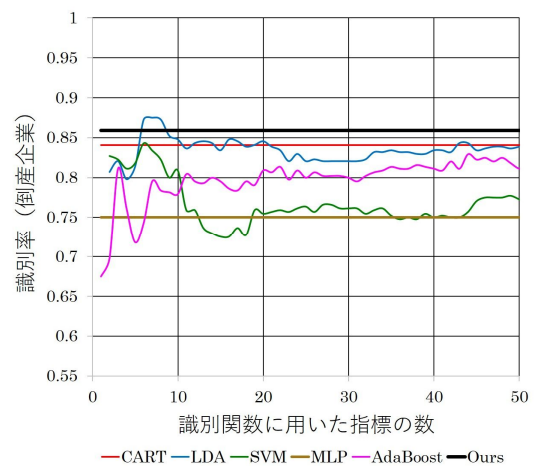


図4：比較手法の性能(実質破綻企業に対する識別率)、Oursは提案手法を示している。

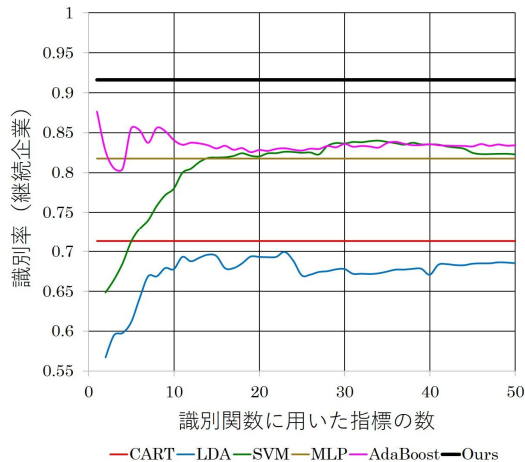


図5：比較手法の性能（継続企業に対する識別率），Oursは提案手法を示している。

## 5. 主な発表論文等

（研究代表者，研究分担者及び連携研究者には下線）

〔雑誌論文〕（計3件）

Yuta Takata, Tadaaki Hosaka, and Hiroshi Ohnuma,  
 "Boosting Approach to Early Bankruptcy Prediction from Multiple-year Financial Statements,"  
 Asia Pacific Journal of Advanced Business and Social Studies, Volume 3, Issue 2, pp.66-76 (July 2017).  
 DOI: 10.25275/apjabssv3i2bus7

Tadaaki Hosaka and Yuta Takata,  
 "Corporate Bankruptcy Forecast Using RealAdaBoost,"  
 INFORMATION-An International Interdisciplinary Journal, Volume 19, No.6(B), pp.2285-2298 (June 2016).  
<https://ci.nii.ac.jp/naid/40020905300>

保坂忠明，高田悠太，大沼宏，  
 “AdaBoostを用いた実質破綻予測モデルの構築と財務指標選択”，  
 日本経営分析学会年報「経営分析研究」，第32号，pp.29-43（2016年3月）.  
<https://ci.nii.ac.jp/naid/40020816280>

〔学会発表〕（計9件）

保坂忠明  
 「財務比率の画像化と深層学習による企業倒産予測」，  
 経営情報学会 2018年春季全国研究発表大会（2018年3月9日 筑波大学東京キャンパス）.

## 保坂忠明

「深層学習による企業の実質破綻予測」，  
 電子情報通信学会パターン認識・メディア理解(PRMU)研究会（2018年2月19日，和歌山大学）.

## 保坂忠明，高田悠太

「決算報告書に基づくRealAdaBoostによる早期の倒産予測」，  
 2017年電子情報通信学会総合大会(2017年3月22日，名城大学天白キャンパス).

Yuta Takata, Tadaaki Hosaka, and Hiroshi Ohnuma,  
 "Boosting Approach to Early Bankruptcy Prediction from Multiple-year Financial Statements,"  
 4th Asia Pacific Conference on Advanced Research, (Hotel Grand Chancellor, Melbourne, Australia, March 2017).

## 高田悠太，保坂忠明，大沼宏

「実質破綻の早期発見を目指したブースティングによる予測モデルの構築」，  
 2017年経営情報学会 春季全国研究発表大会（2017年3月9日，法政大学市ヶ谷キャンパス）.

## 高田悠太，保坂忠明，大沼宏

「時系列財務指標とAdaBoostを用いた実質破綻予測モデルの構築」，  
 2016年日本経営分析学会年次大会(2016年5月21日，中部大学春日井キャンパス）.

## 保坂忠明，高田悠太

「RealAdaBoostを用いた倒産予測指標の抽出」，  
 2016年電子情報通信学会総合大会(2016年3月17日，九州大学伊都キャンパス）.

Yuta Takata, Tadaaki Hosaka, and Hiroshi Ohnuma,  
 "Financial Ratios Extraction Using AdaBoost for Delisting Prediction,"  
 The Seventh International Conference on Information, (Taipei, Taiwan, November, 2015).

高田悠太，保坂忠明，大沼宏  
「上場廃止企業を判別するための AdaBoost  
を用いた財務指標選択」，  
2015 年日本経営分析学会年次大会( 2015 年 5  
月 16 日 産業能率大学自由が丘キャンパス) .

6 . 研究組織

(1) 研究代表者

保坂 忠明 ( HOSAKA TADAAKI )  
東京理科大学・経営学部・講師  
研究者番号 : 60516235