

Department of Policy and Planning Sciences

Discussion Paper Series

No.1385

リレーションシップバンキング機能の向上を目的とした

中小企業の資金ニーズ判別法とその活用の提案

(Proposal of a Method to Determine the Financial Needs of SMEs and Its
Utilization for the Purpose of Improving Relationship Banking Functions)

by

樽 涼稀, 竹田俊彦, 今倉 暁, 櫻井鉄也, 岡田幸彦

(Ryoki MOTAI, Toshihiko TAKEDA, Akira IMAKURA,

Tetsuya SAKURAI and Yukihiro OKADA)

March 6, 2023

UNIVERSITY OF TSUKUBA

Tsukuba, Ibaraki 305-8573

JAPAN

リレーションシップバンキング機能の向上を目的とした

中小企業の資金ニーズ判別法とその活用の提案

Proposal of a Method to Determine the Financial Needs of SMEs and Its Utilization for the Purpose of Improving Relationship Banking Functions

筑波大学大学院サービス工学学位プログラム 罇 涼稀

水戸信用金庫 竹田俊彦

筑波大学システム情報系／人工知能科学センター 今倉 暁

筑波大学システム情報系／人工知能科学センター 櫻井鉄也

筑波大学システム情報系／人工知能科学センター 岡田幸彦

〈論文要旨〉

リレーションシップバンキングは、金融機関が顧客と親密な関係を築き、長期的な取引を行うことで収益を得るビジネスモデルを指す。中小・地域金融機関はリレーションシップバンキング機能を有しているため、財務上の制約を受ける中小企業を支える社会基盤としての役割が期待されている。しかし、中小・地域金融機関は、取引先の負担を過度に引き受けることにより、本来のリレーションシップバンキングの在り方から乖離している面が見られる。また、リレーションシップバンキングは、個々の取引の採算を重視するトランザクションバンキングとは対照的に、近年急速に発展したIT技術を十分に活用することができていない。そのため、IT技術を用いたリレーションシップバンキング業務の効率化が求められている。本研究では、これらの課題を解決するため、企業が特定できないよう秘匿化加工した状態で信用金庫内のみでデータ解析を行うことによって、中小企業の資金ニーズを80%以上の精度で判別するモデルを開発した。本研究の学術的な貢献は、資金ニーズを判別する問題設定について、既存研究を上回る性能を示すモデルを開発したことである。社会的な貢献は、開発したモデルを、実際のリレーションシップバンキングとしての業務に応用する方法を示したことである。本研究で提案した手法は、一定量の財務諸表データを有する金融機関であれば、実行可能である。

〈キーワード〉

中小・地域金融機関, リレーションシップバンキング, 長期借入金, 機械学習, アカウンティング・インフォマティクス

1 問題設定

1.1 社会的背景

金融機関が取り扱う業務は、預金・融資・為替を中心として多岐にわたり、活動目的や顧客、営業エリアに応じた各々のビジネスモデルを展開している。その中でも、一般企業や個人と取引を行う金融機関が取り扱うビジネスモデルは、トランザクションバンキングとリレーションシップバンキングに分類することができる。金融審議会金融分科会第二部会が公表したリレーションシップバンキングの機能強化に向けて（平成15年3月）¹では、トランザクションバンキングとは、「個々の取引ごとの採算性を重視する銀行経営手法であり、貸出に当たっては財務諸表や客観的に算出されるクレジットスコアといった定量的な指標を重視する」（金融審議会金融分科会第二部会 2003, p.5）ビジネスモデルとされている。一方、リレーションシップバンキングとは、「金融機関が顧客との間で親密な関係を長く維持することにより顧客に関する情報を蓄積し、この情報を基に貸出等の金融サービスの提供を行うことで展開するビジネスモデルを指すのが一般的」（金融審議会金融分科会第二部会 2003, p.3）とされている。

これら2つのビジネスモデルについて、滝川（2007）は、大銀行（都市銀行・信託銀行）は、トランザクションバンキング志向の立場を取り、中小・地域金融機関（地方銀行・第二地方銀行・信用金庫・信用組合）は、リレーションシップバンキング志向の立場を取るとした²。また、金融庁が公表した中小・地域金融機関向けの総合的な監督指針（令和4年8月）³においても、中小・地域金融機関は、リレーションシップバンキング機能を有しており、主要な融資対象は中小企業や個人であるとされている。さらに、金融審議会金融分科会第二部会（2003）では、信用金庫や信用組合などの非営利組織である協同組織金融機関には、リレーションシップバンキングを担う存在となることが期待されている。

中小企業の資金調達に関して、中小企業は資本市場で財務的な制約を受けるため、大企業と異なる行動を取ると考えられている（Gracia and Mira 2008）。実際に、中小企業庁が公表した2021年度版中小企業白書⁴からは、我が国中小企業の借入金依存度⁵は大企業と比較して高い傾向にあり、小規模企業において、その傾向は顕著であることがわかる。ゆえに、中小企業は、中小・地域金融機関と良好な関係を維持し、必要な時に必要な融資を受けられる状態にあることが望ましい。

しかし、金融庁の中小・地域金融機関向けの総合的な監督指針（令和4年8月）では「中小・地域金融機関は、地方経済を取り巻く厳しい環境の下、中小企業や地域経済から期待される役割を果たすため、取引先や地域への過大なコミットメントコスト⁶を負担することにより、かえって収益力や健全性の低下といった状

¹ <https://www.fsa.go.jp/news/newsj/14/singi/f-20030327-1.pdf>

² 同著は Berlin and Mester（1998）を引用し、トランザクションバンキングとリレーションシップバンキングは、必ずしも二項対立的な概念ではないことも述べている。

³ <https://www.fsa.go.jp/common/law/guide/chusho/index.html>

⁴ <https://www.chusho.meti.go.jp/pamflet/hakusyo/2021/chusho/index.html>

⁵ 借入金依存度は「（金融機関借入金+その他の借入金+社債）/総資産」で計算される。中規模企業は資本金1,000万円以上1億円未満の企業であり、小規模企業は資本金1,000万円未満の企業とされている。

⁶ 金融審議会金融分科会第二部会（2003）では、コミットメントコストとして、次の3点を取り上げられている。「①金利水準からは正当

況を招いている場合がある。このように、中小・地域金融機関の実態は、リレーションシップバンキング本来のあり方から乖離している面があり、リレーションシップバンキングの機能強化を図り、地域の金融ニーズへの一層適切な対応や、持続可能性（サステナビリティ）の確保を図る必要があると考えられる」（金融庁 2023, 注釈著者）とされている。したがって、中小・地域金融機関は、リレーションシップバンキング機能の強化を図るために、限られた営業エリア内の資金ニーズを読み解き、収益力と健全性の維持・向上に努めることが求められている。

また、近年の IT 技術の発展は、金融機関の業務を効率化することで、収益力の向上に寄与している。特に、トランザクションバンキングにおいては、IT 技術との親和性が高いために、IT 技術を用いた業務の効率化が進められている（Jakšič and Marinč 2017）。その一方で、リレーションシップバンキングへの IT 技術の適用は、トランザクションバンキングと比較して遅れている（Jakšič and Marinč 2017）。したがって、IT 技術を用いたリレーションシップバンキング関連業務の効率化を図るための方法論が求められている。

1.2 学術的背景

学術研究において、トランザクションバンキングとリレーションシップバンキングについては、以下の定義が与えられている（Boot 2000）。本研究では、Boot（2000）の定義に基づいたリレーションシップバンキングを研究の対象とする。

トランザクションバンキング：一回の取引で収益性を評価する。

リレーションシップバンキング：同一の顧客と長期的に（または複数の）商品を取引する中で、顧客が保有する私的な情報を得て、総合的な収益性を評価する。

次に、リレーションシップバンキングと中小企業の関係に注目した研究の一部を取り上げる。Berger and Udell（1995）や Elyasiani and Goldberg（2004）は、中小企業が金融機関に借入を申し込む際に情報の非対称性が問題となることを指摘し、金融機関がリレーションシップバンキング機能を有することで、情報の非対称性を緩和することが可能であると指摘した。また、Jakšič and Marinč（2017）は、リレーションシップバンキングに関する文献レビューを行い、リレーションシップバンキング機能を有する金融機関の存在は、金融市場から資金調達を行うことが困難な中小企業にとって特に価値があるため、安定な社会経済の基盤となることが期待されていると結論付けた。

金融機関は、企業の特徴を把握するために、財務諸表から作成される財務比率を参照する⁷。そこで、財務諸表を用いて、企業の将来を定量的に予測した研究に注目する。Altman（1968）による倒産予測や、Ou and Penman（1989）による収益予測は、財務諸表から企業活動の結果を予測した古典的な研究である。そして、これらの研究を嚆矢として、企業の財務情報に機械学習や深層学習の知見を応用し、企業活動の結果を予測

化できない信用リスクの負担」、②「地域における悪評の発生（レピュテーションリスク）を恐れた問題の先送り」、③「採算性を離れたサービスの提供」。

⁷ 水戸信用金庫におけるヒアリングより。例えば、伝統的な財務諸表分析における収益性指標や安全性指標などである。

した研究が数多く報告されている (Nosratabadi et al. 2020)⁸。会計学分野においても、AI 技術を応用した研究が蓄積されているが、銀行が活用している AI 技術に関する学術研究は、ほとんど存在しないことが指摘されている (首藤 2019)。なお、2023 年現在では、銀行が AI 技術を用いて、信用審査の自動化や業績予測を行うことで、業務を効率化した事例が報告され始めている⁹。しかし、それらの実用可能性を示す学術研究は、ほとんど報告されていないままである。

そのような中、岡田ほか (2021) と罇ほか (2022) は、リレーションシップバンキング機能を強化するために、企業の資金ニーズを判別する機械学習モデルを構築し、その方法論を提示した。しかし、それら 2 つの研究が使用したデータは、東証一部上場企業の財務諸表データであった。リレーションシップバンキング機能を持つ金融機関の主要取引先は中小企業であるため (金融庁 2021)、それら 2 つの研究で提示されたモデル構築のための方法論が、中小企業に対しても適用可能かを明らかにする必要がある。また、それら 2 つの研究では、構築した機械学習モデルの結果を金融機関がどのように用いるかは明言されていない。

なお、会計学やファイナンスの分野では、財務情報と借入金の関係について、統計分析を用いて明らかにした研究が存在する (Chittenden et al. 1996; Michaelas et al. 1999; Öhman and Yazdanfar 2017 など) が、中小企業の財務諸表データを用いて、資金ニーズを判別した研究は見当たらない。

1.3 本研究の目的

本研究の目的は、金融庁 (2023) で挙げられた課題の 1 つである「地域の金融ニーズへの一層適切な対応」を、実際の中小・地域金融機関で行うための方法論を提案することである。そこで本研究では、金融庁 (2023) が指摘した課題を、①金融機関が取引先の資金ニーズを事前に察知すること、②資金ニーズを抱えている取引先へ主体的に訪問・営業を行うことの二段階に分解し、中小・地域金融機関が各段階を実行するための方法論を提案する。

まず初めに、本研究で扱う資金ニーズを設定する。序列理論 (Pecking Order Theory) によると、中小企業が資金を必要とした際に、資金の調達先として考える順位は、内部留保、短期借入金、長期借入金、株式の順である (Myers 1984; Chittenden et al. 1996; Öhman and Yazdanfar 2017; Boateng et al. 2022)。したがって、貸借対照表の長期借入金の変動には、内部留保や短期借入金で賄うことができなかった、明らかな資金ニーズが反映されている。さらに、「長期借入金は地域金融機関にとっては収益力の源泉」(罇ほか 2022, p.2) であるため、長期借入金の変動を判別することは、中小・地域金融機関の実務においても有用な行為である。以上の事柄を踏まえ、本研究では、企業の長期借入金の増加には、当該企業の資金ニーズが反映されていると仮定する。そして、中小企業の資金ニーズを察知するために、長期借入金の変動を判別する数理モデルを

⁸ 経済学分野において、機械学習や深層学習を用いた 57 件の論文をレビューしたものである。その内訳は、株式市場に関する研究 37 件、マーケティング 6 件、電子商取引 8 件、企業倒産 3 件、暗号通貨 3 件であった。すべての論文が財務諸表データを取り扱うわけではなく、非財務情報を利用した研究も紹介されていた。

⁹ 日本経済新聞 2020 年 12 月 8 日朝刊など。

構築することとする。

罇ほか（2022）では、「リレーションシップバンキングの機能強化を図り、地域の金融ニーズへの一層適切な対応や、持続可能性の確保を図るには、どの地域金融機関でも有する取引先の財務諸表を活かし、機械学習を応用した費用対効果の高い地域の資金ニーズの察知が重要となる」（罇ほか 2022, p.2）という目的意識のもとで、財務諸表から長期借入金の増減を判別するための方法論が提示された。これは、リレーションシップバンキング機能を有する中小・地域金融機関において、機械学習を用いて長期借入金の増減を判別する数理モデルを導入し、実際に収益力と健全性の維持・向上がなされるまでの到達段階を考えたとき、TRL2 から TRL3（TRL：Technology Readiness Level, 表 1 参照）に相当する。

表 1 TRL の段階

TRL 9	運用環境内でシステムを実証
TRL 8	システムの完成と認証
TRL 7	運用環境内で行う試作機のデモンストレーション
TRL 6	関連環境内で行う技術のデモンストレーション
TRL 5	関連環境内で技術を検証
TRL 4	ラボ環境内で技術を検証
TRL 3	コンセプトの実験的証明
TRL 2	技術コンセプトの策定
TRL 1	基本原理の観測

（出典）NASA Technology Readiness Level をもとに筆者作成。

本研究は、罇ほか（2022）の目的意識を引き継ぎ、実際の中小・地域金融機関において、①金融機関が取引先の資金ニーズを事前に察知すること、および②資金ニーズを抱えている取引先へ主体的に訪問・営業を行うための方法論を提案する。すなわち、①について、中小・地域金融機関の実際の取引先企業の財務諸表データを用いて、長期借入金の増減を判別する数理モデルを構築する（TRL4 から TRL5）。本研究は、水戸信用金庫との共同研究¹⁰であり、共同研究契約および秘密保持契約を厳守して実施された。水戸信用金庫が保有する取引先事業者の財務諸表データは、取引先事業者の機密データである。そのため、当該データについて企業が特定できないかたちで秘匿化加工を行い、かつ、秘匿化加工された分析用データは外部に持ち出さず、水戸信用金庫本店内の定められた部屋でのみデータ解析とモデルの構築を行った（使用データの詳細は 3 節 1 項を参照）。次に、②について、構築したモデルが出力した判別結果を営業業務に利用する際の例を示す。これは、判別結果に基づく施策が、中小・地域金融機関における収益力と健全性の維持・向上にどの程度貢献したのか検証する TRL5 から TRL6 への足掛かりを作成する試みである。

¹⁰ 水戸信用金庫は、信用金庫法に定められた非営利の団体である。また、Okada and Inamizu（2014）のように、大学における教育研究に積極的に寄与してきた歴史を有する。

2 先行研究

2.1 財務諸表から企業活動の結果を予測した研究

財務諸表を用いて、企業の収益増加、あるいは減少を予測した古典的な研究として、Ou and Penman (1989) が挙げられる。Ou and Penman (1989) は、財務諸表分析により、株価に反映されていない本質的な企業価値を発見する方法を提案した。それは、複数企業の財務諸表から、当該企業の収益が次年度に増加する確率値を予測するロジスティック回帰モデルを構築し、予測結果に基づく投資戦略をとり、投資のリターンを分析することで達成された。Ou and Penman (1989) は、作成したロジスティック回帰モデルの予測精度を、モデルの学習に使用していないテストデータを用いて測定しており、機械学習の手法を用いた研究を行った。Ou and Penman (1989) の分析により、財務諸表には少なくとも企業の収益の方向性（増えるか否か）を予測する能力があることが明らかとなった。これは、財務諸表から長期借入金の増減を判別する数理モデルの開発を後押しするものである。

Altman (1968) は、会計学分野において伝統的に重視されていた単変量の財務比率分析を、多変量に拡張することを目的の 1 つとして、5 つの財務比率から企業の倒産を判別する多変量判別モデル (式 1) を開発した。

$$Z = 0.012X_1 + 0.014X_2 + 0.033X_3 + 0.006X_4 + 0.999X_5 \quad (1)$$

ここで、 X_1 は「運転資本/総資産」、 X_2 は「利益剰余金/総資産」、 X_3 は「支払利息・税引前当期純利益/総資産」、 X_4 は「株主資本の市場価値/総負債」、 X_5 は「売上高/総資産」である。Altman (1968) は、 Z 値が 2.99 以上の企業は非倒産群に属し、 Z 値が 1.81 以下の企業は倒産群に属すると説明した。また、 Z 値が 1.81 から 2.99 の間に属する企業については、明確に倒産の有無を分けることができないが、閾値を 2.675 とすることを提案した。Altman (1968) が開発した多変量判別モデルは、1 年後から 2 年後までの倒産であれば、70% 以上の判別性能を発揮した。しかし、Altman (1968) の多変量判別モデルでは、3 年後から 5 年後までに倒産する企業を予測することは困難であるという実験結果が得られた。これらの結果から、Altman (1968) は、自らが構築した多変量判別モデルには、未来予測の観点から限界があることを認識したうえで、モデルの出力結果は、銀行員が融資を行う際の判断材料として有用であることを示唆した。

Perboli and Arabnezhad (2021) は、中小企業の財務情報と定性的な非財務情報を使用して、1 年後から 5 年後までの倒産を 90% 以上の精度で予測するモデルを開発することで、Altman (1968) の多変量判別モデルが抱えていた、5 年後までに倒産する企業を予測するという課題を解決した。加えて、Perboli and Arabnezhad (2021) は、開発した倒産予測モデルを用いて、COVID-19 が経済へ与えた影響の分析を行った。彼らが開発したモデルは、ARISK 社が提供する意思決定支援システムに組み込まれている (Perboli and Arabnezhad 2021)。

このように、倒産予測の分野では、倒産に関連する財務比率が個別に複数特定され、多変量を扱う倒産予測モデルが構築されることで、倒産予測モデルの精度向上と予測可能期間の拡大を目的とした研究が蓄積されてきた。加えて、構築された倒産予測モデルは、政策立案や商用利用のために応用され始めている。一方、

本研究の対象である資金ニーズ、あるいは長期借入金の増減を判別・予測する研究は、十分に蓄積されているとは言い難い。財務諸表の情報から資金ニーズを判別した学術研究としては、次節で紹介する罇ほか(2022)が代表的であると思われる。

2.2 財務諸表から長期借入金の増減を判別した研究

始めに、長期借入金に関連する財務比率を整理・実証した研究として、Öhman and Yazdanfar (2017) が挙げられる。Öhman and Yazdanfar (2017) は、スウェーデンの中小企業において、企業規模が大きいほど長期借入金を利用する可能性が低いこと、成長性が高いほど長期借入金を利用する可能性が高いこと、収益性や流動性が高い企業ほど長期借入金を利用する可能性が低いことを、統計分析によって明らかにした。Öhman and Yazdanfar (2017) が明らかにした、長期借入金と統計的に有意な連関を持つ財務比率は、本研究においても判別性能の向上に寄与することが期待される。ただし、Öhman and Yazdanfar (2017) の分析は、 t 期の負債を t 期の状態から回帰したものである。すなわち、Öhman and Yazdanfar (2017) は、実際の企業が有する資金ニーズを判別するための数理モデル開発を行っていないことに注意する必要がある。

罇ほか(2022)は、NEEDS 財務データ¹¹より取得した東証一部上場企業の $t-1$ 年度と t 年度の財務諸表データを使用し、 t 年度から $t+1$ 年度にかけての当該企業の長期借入金増減を判別する方法論を提示した($t = 2005, \dots, 2015$)。つまり、 t 年度から $t+1$ 年度にかけて長期借入金が増加した企業と、減少した企業を分類する判別モデルを構築した。判別対象となる目的変数は、「 $t+1$ 年度の長期借入金/ t 年度の長期借入金」が1以上である企業を1として、1未満である企業を0とした二値の変数である。サンプル数は24,471であり、説明変数の候補は、 $t-1$ 年度の財務諸表と t 年度の財務諸表から作成可能な2,559個であった(表2参照)。

表2 モデル構築に使用した説明変数の候補(罇ほか 2022)

説明変数の種類	説明変数の略称	個数
$t-1$ 年度の財務諸表の値	$t-1$ 年度の値	425
t 年度の財務諸表の値	t 年度の値	425
(t 年度の値) - ($t-1$ 年度の値)	値の差	425
$t-1$ 年度の財務諸表の値 / $t-1$ 年度の総資産額	$t-1$ 年度の比	424
t 年度の財務諸表の値 / t 年度の総資産額	t 年度の比	424
(t 年度の比) - ($t-1$ 年度の比)	比の差	424
$t-1$ 年度から t 年度にかけての売上増減ダミー変数	売上増減ダミー	1
年度のダミー変数	年度ダミー	11

(出典) 罇ほか(2022)をもとに筆者作成。

¹¹ <https://www.nikkei.co.jp/needs/services/financial-quest> (契約者のみアクセス可能)

罇ほか（2022）では、上記の説明変数候補と目的変数を含むデータに対して前処理を行った後、長期借入金の増減を判別するモデル構築のための方法論が探索された。なお、前処理後データのサンプル数は12,988であり、421個の説明変数を持っていた。方法論の探索は、四段階に分けて行われた。一段階目は、年度によるデータの事前分割の有無である。年度によるデータ分割を行う場合、基準年度ごとに判別モデルが構築されることを意味する。二段階目は、データのスケールリングについてであり、【標準化・正規化】のいずれかがデータに適用された。三段階目は次元削減処理であり、【PCA（Principal Component Analysis, 主成分分析）・SVD（Singular Value Decomposition, 特異値分解）・FA（Factor Analysis, 因子分析）・ICA（Independent Component Analysis, 独立成分分析）】のいずれかを用いて、 n 次元にデータが圧縮された（ $n = 10, 50, 100, 200, 300, 400, 410$ ）。四段階目は変数選択であり、目的変数との相関が強い説明変数が m 個選択された（ $m = 10, 50, 100, 200, 300, 400, 410, m \leq n$ ）。なお、標準化や次元削減、変数選択の各処理を行わない場合も含めて探索が行われた。判別に使用したモデルは、損失関数にL2正則化項を加えたリッジ・ロジスティック回帰モデルであった。判別性能の評価には、正解率を表すACC（Accuracy）と、真陽性率と偽陽性率から図示されるROC曲線の面積を表すAUCが用いられた（2つの指標の説明は3節5項を参照）。

前述の方法論探索の結果、基準年度 t ごとにデータを分割し、各分割データを標準化し、SVDを用いて410次元に圧縮した後に、目的変数との相関が強い説明変数を100個選択したうえで、モデルの学習を行うこと（方法論1）が、最も高い判別性能を示す結果が得られた（ACCの平均値は0.730、AUCの平均値は0.757）。

次に、罇ほか（2022）では、水戸信用金庫とのヒアリングを通じて得た「財務的に苦しいからお金を借りるというようなネガティブな理由の融資案件と、経営はうまくいっており、リニューアルや事業拡大のためにお金を借りるというようなポジティブな理由の融資案件がある」（罇ほか 2022, p.14）というドメイン知識が応用された。すなわち、 $t-1$ 年度から t 年度にかけて、売上が増加した企業と減少した企業にあらかじめデータを分割したうえで、方法論1が適用された。これにより、ACCの平均値は0.801となった。年度や売上によるデータ分割前の目的変数の分布から、ACCの期待値は0.550であることがわかるため、全体として+25.1%ポイント程度の判別性能の向上が見られたことになる。これを踏まえ、罇ほか（2022）では、表3の方法論が提案された。表3の方法論を、本研究では方法論2と呼称し、中小企業の財務諸表データへの適用を試みる。

表3 方法論2

1	各基準年度 t でデータを分割する
2	$t-1$ 年度から t 年度にかけて、売上が増加した企業と減少した企業でデータを分割する
3	各分割データを標準化する
4	標準化後の各分割データをSVDで410次元に圧縮する
5	圧縮後の各分割データの説明変数について、目的変数との相関が強いものを100個選択する

3 研究方法

3.1 使用データ

本研究で用いるデータは、共同研究先である水戸信用金庫が保有する取引先事業者の、企業が特定できないように秘匿化加工された財務諸表データである。期間は2016年1月から2021年3月までであり、のべ27,068事業者の財務諸表データとなっている。

信用金庫の会員は、信用金庫法第二章第十条¹²により、「一 その信用金庫の地区内に住所又は居所を有する者」「二 その信用金庫の地区内に事業所を有する者」「三 その信用金庫の地区内において勤労に従事する者」「四 前三号に掲げる者に準ずる者として内閣府令で定める者」に限定されている。加えて、「法人にあつてはその常時使用する従業員の数が三百人を超え、かつ、その資本金の額又は出資の総額が政令で定める金額を超える事業者を除くもの」とされている。一般社団法人全国信用金庫協会（2021）によると、法令で定める金額については、9億円以下とされている。本研究で使用するデータについて、事業者規模の分布を確認すると、資本金9億円以上のデータが25サンプル存在していた（全データ中0.09%）。これらの事業者は、卒業生金融¹³の対象と考えられるため、以降の分析から除外した。

また、本研究の目的の1つは、リレーションシップバンキング機能を有する信用金庫を対象として、取引先事業者の長期借入金の増減を判別するモデルを構築することである。そのため、中小企業基本法が定義する中小企業ではなく、信用金庫が通常取引可能な事業者の財務諸表データを用いて、以降の分析を進めることとした¹⁴。使用データに含まれる年度ごとの事業者数と、説明変数として利用可能なものを表4に示す。ただし、 t 年度とは、 t 年4月1日から $t+1$ 年3月31日までの期間を意味する（例えば、2015年度は、2015年4月から2016年3月まで）。そのため、2015年度の事業者数は、他の年度と比較して少ない。2015年度を除けば、単年度あたりの平均事業者数は、5,149サンプルである。

表4 年度ごとの事業者数と説明変数の内訳

年度	事業者数	分類	個数
2015	1,299	財務諸表項目	194
2016	4,904	財務比率	87
2017	5,332	経常収支比率関連	22
2018	5,476	変動計算書	36
2019	5,283	脚注（従業員数等）	22
2020	4,749	その他（顧客番号等）	15
全期間	27,043	合計	376

¹² <https://elaws.e-gov.go.jp/document?lawid=326AC1000000238>

¹³ 信用金庫法施行令第八条によると、事業規模の拡大等により会員資格を失った事業者に対して、一定の制限のもと貸付が許可されている。

¹⁴ 中小企業基本法第二条で定義される中小企業に該当しない事業者は、114サンプルと僅かであった（資本金9億円以上の事業者を取り除いたデータ中0.42%）。

本研究では、中小・地域金融機関において、電子データとして保管されている可能性が最も高く、信頼性の高い情報である財務諸表データ（罇ほか 2022）の利用可能性に焦点を当てる。表 4 で示した財務諸表項目以外の説明変数候補については、以下の理由で取り扱わないこととした。

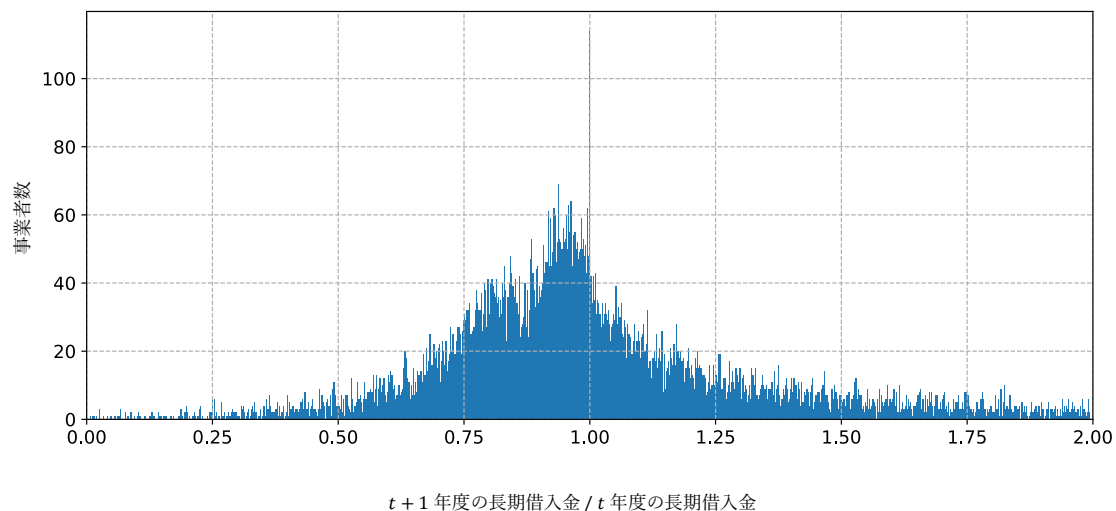
- 財務比率：概ね財務諸表項目から計算可能であるが、金融機関によっては、計算している財務比率の種類や計算式が異なる可能性があるため。
- 変動計算書：欠損値が多く、一部の事業者のみで使用されているため。
- 非財務情報：従業員数等の情報は、金融機関によっては集計していない可能性があるため。加えて、財務諸表データが長期借入金の増減をどの程度判別する能力を有するかを検証する意図があるため。

したがって、以降の分析では、財務諸表項目 194 個（内訳は、貸借対照表項目 95 個、損益計算書項目 59 個、キャッシュフロー計算書項目 40 個）を説明変数の候補とした。

3.2 前処理

本研究の目的の 1 つは、中小・地域金融機関の秘匿化加工された取引先事業者の財務諸表データを用いて、当該事業者の長期借入金の増減を判別することである。したがって、目的変数は、「 $t+1$ 年度の長期借入金 / t 年度の長期借入金」が 1 以上である事業者を 1 として、1 未満である事業者を 0 とした二値の変数とした（罇ほか 2022）。これにより、罇ほか（2022）における結果と比較することが可能となる。なお、分母が 0 の事業者は、差分を用いて目的変数を設定しており、長期借入金の値が欠損している事業者については、目的変数を作成することができないため、以降の分析から除外した。参考として、「 $t+1$ 年度の長期借入金 / t 年度の長期借入金」を二値に変換する前の分布を図 1 に示す。目的変数を「 $t+1$ 年度の長期借入金額」や「 $t+1$ 年度の長期借入金 / t 年度の長期借入金」の値にすることで、値自体の予測を行う回帰モデルの構築も可能であるが、外れ値に該当する目的変数を持つ事業者が過大評価される可能性を考慮した（Ou and Penman 1989）。

図 1 「 $t+1$ 年度の長期借入金 / t 年度の長期借入金」の分布



0 から 2 の間を 1,000 分割したヒストグラムである。基本統計量は、最小値 0.007, 25%点 0.829, 中央値 0.977, 75%点 1.239,

最大値 491.650, 平均値 1.472, 標準偏差 6.883 であり、ヒストグラムの全体像は右の裾が長い形状である。

Öhman and Yazdanfar (2017) の研究では、目的変数に当たる長期借入金と、説明変数に当たる企業規模や収益性などは、同年度のデータを用いていた。しかし、本研究は予測を行う視座からモデルを構築するため、目的変数と説明変数の間に 1 年度分のラグを設定した。また、モデルの学習に使用する説明変数の期間は、罇ほか (2022) と同様に、過去 2 年分とした。すなわち、本研究は、 $t-1$ 年度と t 年度の財務諸表データを用いて、 $t+1$ 年度における長期借入金の増減を判別するモデルの構築を行う。そのため、モデルの構築には $t-1$ 年度から $t+1$ 年度にかけての、連続した 3 年度分の財務諸表データが必要である。したがって、連続した 3 年度分の財務諸表データを持たない事業者を取り除いた。

Öhman and Yazdanfar (2017) の研究では、目的変数に当たる長期借入金と、説明変数に当たる企業規模や収益性などは、同年度のデータを用いていた。しかし、本研究は予測を行う視座からモデルを構築するため、目的変数と説明変数の間に 1 年度分のラグを設定した。また、モデルの学習に使用する説明変数の期間は、罇ほか (2022) と同様に、過去 2 年分とした。すなわち、本研究は、 $t-1$ 年度と t 年度の財務諸表データを用いて、 $t+1$ 年度における長期借入金の増減を判別するモデルの構築を行う。そのため、モデルの構築には $t-1$ 年度から $t+1$ 年度にかけての、連続した 3 年度分の財務諸表データが必要である。したがって、連続した 3 年度分の財務諸表データを持たない事業者を取り除いた。

表 5 モデル構築に使用する説明変数の候補

説明変数の種類	説明変数の略称	個数
$t-1$ 年度の財務諸表の値	$t-1$ 年度の値	194
t 年度の財務諸表の値	t 年度の値	194
$(t$ 年度の値) $-$ ($t-1$ 年度の値)	値の差	194
貸借対照表項目： $t-1$ 年度の財務諸表の値 / $t-1$ 年度の総資産額 ^a それ以外： $t-1$ 年度の財務諸表の値 / $t-1$ 年度の売上高 ^b	$t-1$ 年度の比	193
貸借対照表項目： t 年度の財務諸表の値 / t 年度の総資産額 ^a それ以外： t 年度の財務諸表の値 / t 年度の売上高 ^b	t 年度の比	193
$(t$ 年度の比) $-$ ($t-1$ 年度の比)	比の差	193
$t-1$ 年度から t 年度にかけての売上増減ダミー変数	売上増減ダミー	1
年度のダミー変数 ^c	年度ダミー	4

^a「総資産額/総資産額」は計算しない。^b「売上高/売上高」は計算しない。

^c後述するが、2016 年度から 2019 年度が基準年度 t として設定可能である。

作成した説明変数の合計は 1,166 個であるが、学習に用いるデータの説明変数が多いほど、学習に必要な計算コストが増大する。また、局所最適解といった、学習を困難にする問題の発生を避けるため、あらかじめ

め一定の基準で説明変数を取り除いた。加えて、説明変数に欠損が多いサンプルを取り除いた。データ全体に対して加えた前処理は以下の通りである。

1. 連続した3年度分の財務諸表データを持たないサンプルを削除し、残ったデータを $t-1$ 年度・ t 年度・ $t+1$ 年度が横に並ぶように連結 (27,043 サンプル→14,538 サンプル)
2. 総資産額、売上高のいずれかが欠損、または0であるサンプルを削除
長期借入金額が欠損しているサンプルを削除 (14,538 サンプル→13,058 サンプル)
3. 分散が0である説明変数を削除 (説明変数 1,166 個→説明変数 1,129 個)
4. 欠損率が10%以上の説明変数を削除 (説明変数 1,129 個→説明変数 462 個)
5. 欠損率が10%以上のサンプルを削除 (13,058 サンプル→12,018 サンプル)
6. 欠損値を0で補完

上記の前処理後のデータセットは、1個の目的変数と462個の説明変数を持つ12,018サンプルの財務諸表データとなった。この時点における基準年度 t ごとの事業者数と、説明変数の内訳を表6に示す ($t = 2015, \dots, 2020$)。本研究で作成するモデルは、 $t-1$ 年度、 t 年度、 $t+1$ 年度の財務諸表データが必要である。したがって、基準年度 t として設定可能な年度は、2016年度から2019年度であることに注意されたい (基準年度 t を2015年度、もしくは2020年度に設定した場合、連続した3年度分の財務諸表データを持つ事業者が存在しないことになる)。

表6 基準年度ごとの事業者数と説明変数の内訳 (前処理後)

基準年度 t	事業者数	目的変数が1である 事業者の割合	説明変数の略称	個数
2015	0	-	$t-1$ 年度の値	77
2016	972	0.381	t 年度の値	78
2017	3,593	0.380	値の差	75
2018	3,736	0.373	$t-1$ 年度の比	75
2019	3,717	0.622	t 年度の比	76
2020	0	-	比の差	76
全期間	12,018	0.453	売上増減ダミー	1
			年度ダミー	4

3.3 方法論の適用

本研究では、罇ほか (2022) で提案された方法論2が、主に中小企業で構成される水戸信用金庫の取引先事業者の財務諸表データにおいても適用可能か検証する。方法論2の1番：基準年度によるデータ分割、2番：売上増減によるデータ分割、4番：標準化、5番：目的変数との相関を用いた変数選択については、罇ほか (2022) と同様である。3番：SVDによる次元削減については、説明変数の個数が罇ほか (2022) と異なる

るため、450次元に圧縮することとした。このような設定のもと、判別性能の評価を行う（詳細は3節5項を参照）。

ここで、本研究で用いるSVDによる次元削減に関する数学的な説明を以下に示す。SVDは、任意の m サンプル n 変数の $m \times n$ 行列 \mathbf{X} を $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^t$ の3つの行列に分解する手法である（Stewart 1993）。ただし、 $\mathbf{\Sigma}$ は、非負の特異値 $(\sigma_1, \dots, \sigma_n)$ を対角成分に降順に並べた対角行列であり、 \mathbf{U} と \mathbf{V}^t はそれぞれ、特異値に対応する n 本の特異ベクトルで構成される直交行列である。SVDを用いて次元削減を行う際は、 $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^t$ を式変形して得られる $\mathbf{U} = \mathbf{X}\mathbf{V}^t(\mathbf{\Sigma})^{-1}$ を用いる。つまり、初めに、特異値を大きい順に k 個選択し（ $k \leq n$ ）、対角行列 $\mathbf{\Sigma}_k$ を作る。そして、選択した特異値に対応する \mathbf{V}^t の行ベクトルにより、 \mathbf{V}_k^t を作成する。これらを用いて、 $\tilde{\mathbf{X}} = \mathbf{X}\mathbf{V}_k^t(\mathbf{\Sigma}_k)^{-1}$ 計算することで、 m サンプル k 変数の次元削減後行列 $\tilde{\mathbf{X}}$ が得られる。なお、選択した特異値に対応する \mathbf{U} の列ベクトルを k 本並べた行列を \mathbf{U}_k とすれば、 $\mathbf{X}\mathbf{V}_k^t(\mathbf{\Sigma}_k)^{-1}$ の行列積計算を省略し、 \mathbf{U}_k を直に $\tilde{\mathbf{X}}$ とすることができる。

本研究は、プログラミング言語 Python (Ver 3.8.8) を用いて計算を行った。SVDによる次元削減時には、NumPy (Ver 1.20.1) の `numpy.linalg.svd` 関数を用いており、関数の内部では、LAPACK の `_gesdd` ルーティンを使用している (NumPy 2021)¹⁵。

3.4 判別モデル

3節2項において、目的変数を長期借入金の増減の方向を示す二値変数と設定したが、構築したモデルを実務に応用する際には、0か1の二値分類ではなく、確率値の参照が可能である方が好都合な場合がある。したがって、本研究では、ロジスティック回帰モデルを使用する。ロジスティック回帰モデルは、複数のモデルの分類性能を比較する際に、ベースラインとして使用されている (Byanjankar et al. 2021)。

なお、本研究で用いる説明変数は、財務諸表を基に作成しており、多重共線性を考慮した変数選択を行っていない。このような状況下でロジスティック回帰モデルのパラメータを推定すると、推定結果が不安定になることがある (Cessiet and Houwelingen 1990)。以上の事柄を踏まえ、罇ほか (2022) と同様に、損失関数にL2正則化項を加えたリッジ・ロジスティック回帰モデルを使用することとした。

モデルのパラメータ推定には、scikit-learn (Ver 1.0) で実装されている `sklearn.linear_model.LogisticRegression` を用いており、ソルバーはL-BFGSとした (scikit-learn 2021)。scikit-learnにおいて、リッジ・ロジスティック回帰モデルの損失関数は式(2)で与えられる。ただし、 \mathbf{X} は説明変数の行列、 \mathbf{y} は目的変数のベクトル、 \mathbf{w} は説明変数に関するモデルの係数ベクトル、 c はモデルの切片である ($i = 1, \dots, m$ は、ある1つのサンプルを表す)。なお、式(2)の C はハイパーパラメータであり、本研究ではデフォルト値の1に設定した。損失関数を最小にする \mathbf{w} と c は最尤法で推定され、リッジ・ロジスティック回帰モデルのパラメータとして、

¹⁵ 他に、`sklearn.decomposition.TruncatedSVD`によりSVDを行うことも可能である。しかし、この場合は、特異値に関する情報が次元削減後のデータに反映されないことに注意する必要がある。

判別に使用される。最終的な長期借入金増加確率の出力値は、式 (3) によって与えられる。確率の出力値は 0 か 1 の二値に分類されるが、分類の閾値については 3 節 5 項で説明する。

$$\min_{\mathbf{w}, c} \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^m \log [\exp\{-y_i(\mathbf{X}_i^T \mathbf{w} + c)\} + 1] \quad (2)$$

$$\hat{y} = 1 / \{1 + \exp(-\mathbf{X}^T \mathbf{w} - c)\} \quad (3)$$

3.5 判別性能の評価

どのモデルがより正確に長期借入金の増減を判別したのか判断するための評価指標として、ACC (Accuracy, 正解率) と AUC (Area Under the Curve) を用いる。ロジスティック回帰モデルの出力値 \hat{y} は確率値であるため、分析者は目的に応じた適当な閾値を定めて、閾値を下回ったサンプルを 0、閾値以上のサンプルを 1 と変換することができる。本研究では、ACC の算出に係る閾値の決定のために、Youden Index (Youden 1950) を用いた。Youden Index は (真陽性率-偽陽性率) で求めることができ、本研究では、Youden Index を最大にする閾値を用いて、 \hat{y} から 0、もしくは 1 に変換した。

各サンプルが 0 か 1 に分類されれば、目的変数の実際の値と比較することにより、ACC を算出することができる。ACC は、図 2 で表される混同行列 (Confusion Matrix) の「(真陽性のサンプル数+真陰性のサンプル数) / 全サンプル数」で計算される。本研究では、ACC が ACC の期待値 (以下、Exp ACC) を上回る程度が大きいほど、判別性能が高いモデルが構築されたと判断する。Exp ACC とは、1 つのデータセット (分割後データでも可) に対して、 $\text{Exp ACC} = \max\{\text{目的変数が 1 であるサンプルの割合, 目的変数が 0 であるサンプルの割合}\}$ で定義される。

図 2 混同行列

		目的変数の実際	
		1	0
予測された 分類先	1	真陽性	偽陽性
	0	偽陰性	真陰性

他に、混同行列を用いて、以下の 4 つの指標を計算することができる (Byanjankar et al. 2021 など)。

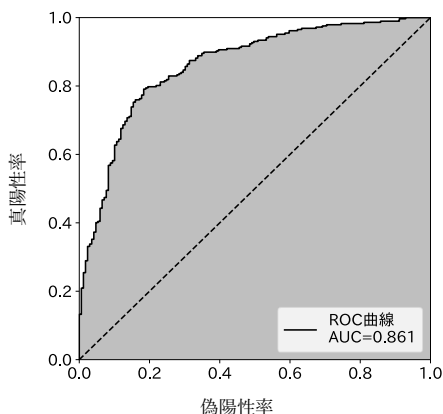
真陽性率 (True Positive Rate) = 真陽性 / (真陽性 + 偽陰性), 偽陽性率 (False Positive Rate) = 偽陽性 / (偽陽性 + 真陰性)

偽陰性率 (False Negative Rate) = 偽陰性 / (真陽性 + 偽陰性), 真陰性率 (True Negative Rate) = 真陰性 / (偽陽性 + 真陰性)

前述のように、ACC の算出には閾値の設定が必要である。これは、同一のモデルであっても、閾値を変化させることで、ACC が変化することを意味している。本研究では、恣意的な閾値の選択による影響を受けない評価指標として、AUC を用いる (Mandrekar 2010)。AUC は、受信者動作特性曲線 (ROC 曲線: Receiver Operating Characteristic Curve) より下部の面積 (図 3 の塗りつぶし部分) であり、二値分類モデルの評価指標として広く使用されている (Lobo et al. 2007)。ROC 曲線は、可能なすべての閾値のもとで (偽陽性率, 真陽性率) の点をプロットし、結ぶことで作成される。仮に、サンプルをランダムに 0 か 1 に分類した場合、ROC

曲線は傾斜 45 度の直線となり、AUC は 0.5 となる (Mandrekar 2010)。したがって AUC が 0.5 を上回る程度が大きいほど、判別性能が高いモデルが構築されたと判断する。

図 3 ROC 曲線と AUC



判別性能を測定するために、データをモデルのパラメータ推定に使用する学習データと、判別性能を測定するためのテストデータに分割する。学習データによるモデルの構築後、再び学習データをモデルに入力し、学習データに対する ACC と AUC (以下、Train ACC と Train AUC) を計算する。同様に、構築されたモデルにテストデータを入力し、テストデータに対する ACC と AUC (以下、Test ACC と Test AUC) を計算する。加えて、評価指標の頑健性を高めるために、層化 20 分割交差検証を行う。層化 k 分割交差検証を行う場合、初めに、目的変数の分布が等しくなるようにデータを k 等分する (D_1, \dots, D_k)。次に、1 つの D_i ($i = 1, \dots, k$) をテストデータ、残りの D_j ($j = 1, \dots, k, i \neq j$) を学習データとし、学習データとテストデータを k 回入れ替えながら、評価指標を k 回計算する。本研究で特に断りなく使用する Train ACC \cdot Train AUC \cdot Train ACC \cdot Train AUC は、層化 20 分割検証時に算出された各指標の平均値である。

参考に、基準年度 t と売上増減によるデータ分割後における、各分割データの事業者数と、目的変数が 1 である事業者の割合、および Exp ACC を表 7 にまとめた。以降は、「 t 年度の売上高 / $t - 1$ 年度の売上高」が 1 以上の事業者を売上事業者とし、1 未満の事業者を売上減少事業者と呼ぶ。

表 7 分割データごとの事業者数と目的変数の割合、および ACC の期待値

A 売上増加事業者				B 売上減少事業者			
基準 年度 t	事業 者数	目的変数が 1 である 事業者の割合	Exp ACC	基準 年度 t	事業 者数	目的変数が 1 である 事業者の割合	Exp ACC
2016	478	0.408	0.592	2016	494	0.354	0.646
2017	1,832	0.403	0.597	2017	1,761	0.355	0.645
2018	2,005	0.397	0.603	2018	1,731	0.345	0.655
2019	1,982	0.637	0.637	2019	1,735	0.605	0.605
全期間	6,297	0.475	0.525	全期間	5,721	0.428	0.572

4 結果

4.1 実験の結果

前処理後データに対して、方法論2の処理（年度や売上増減によるデータ分割、標準化、SVD、目的変数との相関による変数選択）を適用した。その後、リッジ・ロジスティック回帰モデルによる学習と判別を行った。その結果を表8に示す。Thresholdは、確率値を0か1に変換する際の閾値であり、それぞれTrain AUCとTest AUCから算出した¹⁶。

表8を見ると、すべての分割データにおいて、Train ACCとTest ACCはExp ACCを上回っており、Train AUCとTest AUCは0.5を上回っていた。これより、水戸信用金庫の秘匿化加工された取引先事業者の財務諸表データに対して方法論2を適用することで、期待値以上の判別を行うリッジ・ロジスティック回帰モデルが学習されたことが示された。

また、売上増加事業者と売上減少事業者の双方におけるTest AUCの平均値は0.75を上回っており、Test ACCの平均値は0.70を上回っていた。加えて、Test ACCと事業者数をもとに算出した、基準年度 t ごとのACCの加重平均値は、基準年度 t を2016年度に設定した場合から順に、0.878, 0.695, 0.672, 0.699であった（例えば、基準年度 t を2016年度に設定した場合は、 $(0.854 \times 478 + 0.901 \times 494) / (478 + 494) = 0.878$ となる）。これより、方法論2を水戸信用金庫の取引先事業者の財務諸表データに適用した場合、Exp ACCからの判別性能の向上幅は、+15.7%ポイントであった（データ分割前のExp ACCは0.547）。ただし、罇ほか（2022）における実験においては、Exp ACC+25.1%ポイントの判別性能を示すモデルが構築されたことを鑑みると、方法論2は、中小企業の財務諸表データに対しては、効力がやや劣ることが示唆された¹⁷。

次に、Threshold列で示されたYouden Indexにより算出した閾値を見てみると、各分割データについて、Train ThresholdとTest Threshold間の差の平均値は0.007であり、その最大値は0.012であった。注目すべきは、基準年度 t を2019年度に設定した場合のTrain ThresholdとTest Thresholdである。この場合のTrain ThresholdとTest Thresholdは、他の基準年度 t のTrain ThresholdとTest Thresholdよりも高い水準であった（ $t = 2016, \dots, 2018$ ）。これより、一定のACCを保つためには、モデルの学習に使用した変数の種類や分布によって、モデルの出力結果 \hat{y} の解釈を変更することの必要性が示唆された。

最後に、基準年度 t ごとのACCの加重平均値について、基準年度 t を2016年度に設定したときに、最も高い値（0.878）が得られたことに注目する。使用データの都合上、2015年度のデータには、2016年1月から2016年3月までの期間に渡るデータが収録されていた。そのため、基準年度 t を2016年度としたデータに含まれる事業者の大半は、1月末決算から3月末決算の事業者である。これより、データをさらに決算期で分割することで、判別性能が向上する可能性が示唆された。そこで、4節2項では、決算期をもとに四半

¹⁶ 未知のデータを予測する観点からは、Train AUCより算出した閾値を用いて、Test ACCを求めることが望ましい。この点は本研究の今後の課題の1つとする。

¹⁷ 方法論2の真の効力を検証するためには、2017年から2021年までの東証一部上場企業の財務諸表データと、他金融機関が保有する中小企業の財務諸表データが必要であろう。

表 8 分割データごとの判別性能と閾値

A 売上増加事業者

基準年度 t	事業者数	Exp ACC	Train			Test		
			ACC	AUC	Threshold	ACC	AUC	Threshold
2016	478	0.592	0.911	0.968	0.418	0.854	0.897	0.414
			(0.006)	(0.002)	(0.007)	(0.072)	(0.058)	(0.027)
2017	1,832	0.597	0.702	0.766	0.405	0.695	0.711	0.408
			(0.006)	(0.002)	(0.003)	(0.050)	(0.049)	(0.013)
2018	2,005	0.603	0.674	0.746	0.398	0.672	0.692	0.399
			(0.012)	(0.002)	(0.004)	(0.037)	(0.042)	(0.012)
2019	1,982	0.637	0.705	0.764	0.635	0.704	0.715	0.634
			(0.006)	(0.003)	(0.001)	(0.043)	(0.050)	(0.009)
全期間	6,297	0.525	-	-	-	-	-	-
平均値	1,574	0.611	0.710	0.811	0.464	0.703	0.754	0.464

B 売上減少事業者

基準年度 t	事業者数	Exp ACC	Train			Test		
			ACC	AUC	Threshold	ACC	AUC	Threshold
2016	494	0.646	0.932	0.980	0.368	0.901	0.925	0.374
			(0.005)	(0.002)	(0.005)	(0.045)	(0.039)	(0.019)
2017	1,761	0.645	0.697	0.763	0.360	0.696	0.703	0.361
			(0.016)	(0.003)	(0.004)	(0.066)	(0.061)	(0.016)
2018	1,731	0.655	0.695	0.761	0.347	0.672	0.696	0.346
			(0.008)	(0.002)	(0.002)	(0.052)	(0.042)	(0.008)
2019	1,735	0.605	0.697	0.780	0.608	0.694	0.719	0.604
			(0.008)	(0.002)	(0.002)	(0.040)	(0.038)	(0.011)
全期間	5,721	0.573	-	-	-	-	-	-
平均値	1,430	0.636	0.717	0.821	0.421	0.706	0.761	0.421

ACC, AUC, Threshold の各数値は、層化 20 分割交差検証時の平均値であり、括弧内は標準偏差である。

平均値について、Exp ACC と ACC は事業者数を用いた加重平均値であり、AUC と Threshold は列の相加重平均値である。

期単位でデータを分割する処理を適用した後に、方法論 2 の各処理を適用した場合の判別性能の評価を行う。

その準備として、四半期・年度・売上増減でデータを分割した後の事業者数と、目的変数が 1 である事業者の割合、および Exp ACC を表 9 に示す。このとき、 $t-1$ 年度から $t+1$ 年度にかけて決算期を変更した事業

者 (53 サンプル) を取り除いた。四半期は、4 月から 6 月を Q1、7 月から 9 月を Q2、10 月から 12 月を Q3、1 月から 3 月を Q4 と表記した。

表 9 分割データごとの事業者数と目的変数の割合、および ACC の期待値 (四半期分割追加後)

A 売上増加事業者				B 売上減少事業者			
基準年度 t	事業者数	目的変数が 1 である 事業者の割合	Exp ACC	基準年度 t	事業者数	目的変数が 1 である 事業者の割合	Exp ACC
2016-Q4	477	0.407	0.593	2016-Q4	490	0.353	0.647
2017-Q1	521	0.384	0.616	2017-Q1	550	0.342	0.658
2017-Q2	532	0.398	0.602	2017-Q2	529	0.355	0.645
2017-Q3	237	0.388	0.612	2017-Q3	213	0.371	0.629
2017-Q4	541	0.433	0.567	2017-Q4	457	0.359	0.641
2018-Q1	590	0.393	0.607	2018-Q1	507	0.359	0.641
2018-Q2	555	0.409	0.591	2018-Q2	538	0.336	0.664
2018-Q3	277	0.426	0.574	2018-Q3	192	0.313	0.688
2018-Q4	576	0.378	0.622	2018-Q4	483	0.352	0.648
2019-Q1	619	0.504	0.504	2019-Q1	552	0.486	0.514
2019-Q2	637	0.666	0.666	2019-Q2	532	0.635	0.635
2019-Q3	260	0.746	0.746	2019-Q3	249	0.659	0.659
2019-Q4	464	0.713	0.713	2019-Q4	387	0.703	0.703
全期間	6,286	0.475	0.525	全期間	5,679	0.427	0.573

四半期でデータをさらに分割することで、方法論 2 と比較して、長期借入金の増減を判別するモデルの学習に使用可能なサンプル数が減少する。一方、変化するマクロ経済状況や経営環境の変化をより考慮したモデルが学習されることにより、判別性能が向上することも考えられる¹⁸。実際に表 9 では、同一の基準年度 t 内のデータであっても、四半期によって、目的変数が 1 である事業者の割合に差異が見られる。特に、基準年度 t を 2019 年度に設定した場合、他の基準年度とは異なる割合の推移が確認できる。データを四半期単位で分割することにより、このような目的変数、あるいは説明変数の変化を捉えたモデルが学習されることが期待される。

¹⁸ なお、NEEDS 財務データより取得した 2005 年 1 月から 2016 年 12 月の東証一部上場企業 22,751 社のデータを用いる場合、3 月末決算の企業が 16,063 社 (全体の 70.6%) であり、12 月末決算の企業が 2,019 社 (全体の 8.9%) である。そのため、罇ほか (2022) と同様の実験設定に、四半期でデータを分割する処理を加えた場合、十分な学習を行うためのサンプル数が不足する可能性がある。

4.2 追加実験の結果

方法論 2 に四半期でデータを分割する処理を加え、リッジ・ロジスティック回帰モデルによる学習と判別を行った。そこで得た判別性能の評価指標と閾値を表 10 に示す。なお、学習に使用するサンプル数が説明変数の個数を下回る場合、SVD の性質上、次元削減後のデータの次元数は「サンプル数-1」とした。

表 10 より、売上増加事業者と売上減少事業者の双方で、構築されたモデルの Train ACC と Test ACC は、分割後のデータから計算した Exp ACC よりも高い数値を示しており（すべて 0.8 以上）、Test AUC についても 0.8 を上回っていた。Test ACC と事業者数をもとに算出した加重平均値の ACC は、基準年度 t を 2016 年度に設定した場合から順に、0.889, 0.869, 0.864, 0.865 であった。これより、方法論 2 に四半期を用いたデータ分割を加えることで、Exp ACC+32.1%ポイントの判別性能を示すモデルが構築された（データ分割前の Exp ACC は 0.547）。これは、Exp ACC+25.1%ポイントのモデルを構築した罇ほか（2022）を上回るものである。また、Train Threshold と Test Threshold の差の平均値は 0.011, 最大値は 0.046 であった。

以上の実験結果を踏まえ、本研究では、財務諸表データから長期借入金の増減を判別するリッジ・ロジスティック回帰モデルを構築するための方法論を、表 11 の 1 番から 6 番の手順（以下、方法論 3）として提案する。

水戸信用金庫へのヒアリングでは、「概ね 70%ほどの精度で長期借入金の増減を当てることができれば良いだろう」とのことだった。したがって、方法論 3、あるいは方法論 2 に則った長期借入金増減判別モデルの構築を行うことで、本研究の目的の 1 つ目である、中小・地域金融機関の実際の取引先企業の財務諸表データを用いて、長期借入金の増減を判別する数理モデルを構築すること（TRL4 から TRL5）が達成された。また、少なくとも本研究で用いたデータにおいては、罇ほか（2022）で提案された方法論 2 を上回る判別性能を発揮する方法論 3 を得ることができた。

4.3 長期借入金の増減を判別するモデルの利用例

本項では、本研究の目的の 2 つ目である、資金ニーズを抱えている取引先へ主体的に訪問・営業を行うための具体例を示す。詳細な手順はこの後に説明するが、簡潔に言えば、以下の 2 ステップである。

- ① 長期借入金増減判別モデルを構築し、判別した結果、偽陽性となった事業者を特定する。
- ② リレーションシップバンキング機能を有する金融機関は、偽陽性事業者との関係の構築に注力する。

まず初めに、リレーションシップバンキング機能を有する中小・地域金融機関は、データベースに保管されている取引先事業者の財務諸表データを取り出し、3 節 1 項と 3 節 2 項で説明した前処理を行う。次に、前処理後のデータに対して、方法論 3 に則った処理を適用する。そして、学習データとテストデータに分けたうえで、リッジ・ロジスティック回帰モデルを構築する。この時、テストデータを入力して得られる出力結果 \hat{y} と、0 もしくは 1 の分類結果を、事業者名と紐づけた ID と共に保存しておく。この点について、層化 k 分割交差検証を行うことで、すべての事業者の出力結果 \hat{y} と分類先を得ることができる。各事業者の出力結果 \hat{y} と分類先が得られたら、分類先を目的変数の実際の値と照らし合わせることで、偽陽性に該当する事

表 10 分割データごとの判別性能と閾値（四半期分割追加）

A 売上増加事業者

基準年度 t	事業者数	Exp ACC	Train			Test		
			ACC	AUC	Threshold	ACC	AUC	Threshold
2016-Q4	477	0.593	0.924	0.972	0.419	0.877	0.906	0.416
			(0.006)	(0.002)	(0.005)	(0.059)	(0.052)	(0.024)
2017-Q1	521	0.616	0.896	0.965	0.395	0.846	0.883	0.396
			(0.005)	(0.002)	(0.004)	(0.064)	(0.066)	(0.023)
2017-Q2	532	0.602	0.891	0.951	0.409	0.853	0.878	0.406
			(0.006)	(0.003)	(0.004)	(0.073)	(0.070)	(0.015)
2017-Q3	237	0.612	0.998	1.000	0.415	0.983	0.982	0.411
			(0.003)	(0.000)	(0.009)	(0.034)	(0.038)	(0.021)
2017-Q4	541	0.567	0.869	0.935	0.444	0.826	0.839	0.441
			(0.009)	(0.003)	(0.005)	(0.079)	(0.094)	(0.019)
2018-Q1	590	0.607	0.872	0.931	0.411	0.825	0.846	0.409
			(0.009)	(0.003)	(0.007)	(0.072)	(0.074)	(0.022)
2018-Q2	555	0.591	0.882	0.950	0.419	0.848	0.881	0.415
			(0.006)	(0.002)	(0.006)	(0.059)	(0.059)	(0.022)
2018-Q3	277	0.574	0.999	1.000	0.442	0.964	0.975	0.444
			(0.002)	(0.000)	(0.006)	(0.049)	(0.036)	(0.019)
2018-Q4	576	0.622	0.864	0.937	0.388	0.809	0.850	0.383
			(0.007)	(0.003)	(0.004)	(0.062)	(0.057)	(0.020)
2019-Q1	619	0.504	0.856	0.930	0.506	0.827	0.853	0.501
			(0.004)	(0.003)	(0.008)	(0.064)	(0.073)	(0.022)
2019-Q2	637	0.666	0.852	0.932	0.661	0.826	0.853	0.656
			(0.013)	(0.002)	(0.005)	(0.048)	(0.051)	(0.018)
2019-Q3	260	0.746	0.996	1.000	0.723	0.946	0.969	0.742
			(0.003)	(0.000)	(0.007)	(0.062)	(0.040)	(0.014)
2019-Q4	464	0.713	0.949	0.987	0.690	0.903	0.938	0.703
			(0.007)	(0.002)	(0.005)	(0.078)	(0.057)	(0.017)
全期間	6,286	0.525	-	-	-	-	-	-
平均値	484	0.611	0.897	0.961	0.486	0.857	0.896	0.486

B 売上減少事業者

基準年度 t	事業者数	Exp ACC	Train			Test		
			ACC	AUC	Threshold	ACC	AUC	Threshold
2016-Q4	490	0.647	0.920	0.978	0.368	0.902	0.918	0.373
			(0.012)	(0.002)	(0.010)	(0.077)	(0.066)	(0.015)
2017-Q1	550	0.658	0.908	0.968	0.352	0.869	0.886	0.362
			(0.010)	(0.002)	(0.007)	(0.057)	(0.052)	(0.014)
2017-Q2	529	0.645	0.889	0.949	0.374	0.851	0.863	0.371
			(0.011)	(0.002)	(0.009)	(0.048)	(0.061)	(0.021)
2017-Q3	213	0.629	1.000	1.000	0.412	0.968	0.984	0.396
			(0.000)	(0.000)	(0.008)	(0.045)	(0.027)	(0.025)
2017-Q4	457	0.641	0.928	0.973	0.374	0.877	0.894	0.373
			(0.007)	(0.002)	(0.005)	(0.054)	(0.063)	(0.022)
2018-Q1	507	0.607	0.918	0.969	0.372	0.870	0.895	0.374
			(0.006)	(0.002)	(0.003)	(0.060)	(0.060)	(0.020)
2018-Q2	538	0.664	0.914	0.966	0.353	0.857	0.886	0.351
			(0.007)	(0.002)	(0.004)	(0.061)	(0.061)	(0.017)
2018-Q3	192	0.688	1.000	1.000	0.382	0.990	0.993	0.336
			(0.000)	(0.000)	(0.006)	(0.031)	(0.023)	(0.018)
2018-Q4	482	0.648	0.923	0.974	0.369	0.888	0.904	0.368
			(0.008)	(0.002)	(0.007)	(0.063)	(0.087)	(0.017)
2019-Q1	552	0.514	0.864	0.935	0.478	0.819	0.848	0.494
			(0.004)	(0.003)	(0.006)	(0.071)	(0.081)	(0.019)
2019-Q2	531	0.635	0.898	0.959	0.621	0.865	0.876	0.627
			(0.009)	(0.002)	(0.008)	(0.083)	(0.074)	(0.019)
2019-Q3	246	0.658	1.000	1.000	0.637	0.957	0.976	0.666
			(0.000)	(0.000)	(0.007)	(0.070)	(0.042)	(0.016)
2019-Q4	387	0.703	0.972	0.991	0.680	0.897	0.934	0.694
			(0.005)	(0.001)	(0.005)	(0.059)	(0.046)	(0.015)
全期間	5,679	0.573	-	-	-	-	-	-
平均値	437	0.640	0.921	0.974	0.444	0.880	0.912	0.445

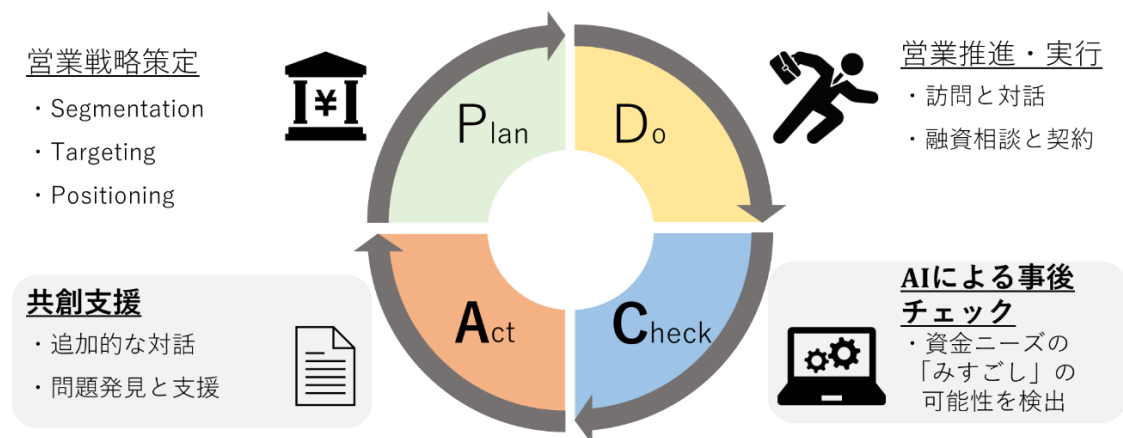
数値の算出法は表 8 と同様である。

表 11 方法論 3

1	各基準年度 t でデータを分割する
2	各四半期でデータを分割する
3	$t-1$ 年度から t 年度にかけて、売上が増加した事業者と減少した事業者でデータを分割する
4	各分割データを標準化する
5	標準化後の各分割データを SVD で 450 次元に圧縮する
6	圧縮後の各分割データの説明変数について、目的変数との相関が強いものを 100 個選択する

業者を特定することができる（図 4 AI による事後チェック部分）。その結果、金融機関の営業担当者は、偽陽性の事業者を重点的に訪問し、それらの事業者と親密な関係を構築することに注力できる（図 4 共創支援部分）。ここで、偽陽性の事業者とは、長期借入金増減判別モデルは長期借入金が増加すると判断したが、実際は増加しなかったものである。そのため、偽陽性の事業者は、偽陽性以外の事業者と比較したとき、資金ニーズを有する可能性が高いと考えられる。

図 4 本研究が提案した方法論に基づく業務フロー



本研究が提案した方法論が、中小・地域金融機関における PDCA サイクルのうち、Check と Act に貢献することを示した図である。

本研究で使用したデータを用いた場合、モデルの出力結果から作成される混同行列は、表 12 で示した結果となった。本研究の場合、873 サンプルが偽陽性に該当した（全体の 7.30%）。最新の基準年度で学習されたモデルの結果のみを用いる場合、201 の事業者が偽陽性に分類された（基準年度 t を 2019 年度に設定した場合の全サンプル中 5.43%）。ただし、閾値の決定に Youden Index を使用しているため、目的に応じた閾値の変更を行うことで、偽陽性の事業者数が増減することが起こり得る。

最後に、偽陽性の事業者を掲載したリストに対する水戸信用金庫の反応を報告する¹⁹。作成した偽陽性事業者のリストには、事業者を示す ID と出力結果 \hat{y} が含まれており、本研究で提案した方法論に基づいて作成したものである。まず総評として、リストに掲載された事業者に違和感はないとのことであった。しか

¹⁹ 水戸信用金庫の担当者が、3 支店を対象に、リストの尤もらしさについて聞き取り調査を行った結果である。

し、決算期に借入を一時的に清算するもの、単体ではなくグループで見ると安定しているもの、無借金経営を理念としているもの、経営内容の問題で融資の対象外とされているものが含まれていた。今後は、最新年度の財務諸表データを用いて作成した偽陽性事業者のリストから、前述の一部事業者を除外したうえで、リストに掲載された事業者へ早速訪問を行う予定であるとのことだった。ただし、本研究に基づいて作成したリストが、実際にどの程度業務の効率化に寄与したのか、あるいは、どの程度収益力の向上に寄与したのかに関する定量的な調査は行っていないことに注意する必要がある。

表 12 本研究の手法を適用して得られた混同行列

A 売上増加事業者 (全 6,286 サンプル)					B 売上減少事業者 (全 5,679 サンプル)				
基準年度 t	真陽性	偽陽性	偽陰性	真陰性	基準年度 t	真陽性	偽陽性	偽陰性	真陰性
2016-Q4	175	40	19	243	2016-Q4	152	27	21	290
2017-Q1	171	51	29	270	2017-Q1	155	39	33	323
2017-Q2	185	51	27	269	2017-Q2	150	41	38	300
2017-Q3	91	3	1	142	2017-Q3	78	6	1	128
2017-Q4	198	58	36	249	2017-Q4	138	30	26	263
2018-Q1	185	56	47	302	2018-Q1	159	43	23	282
2018-Q2	192	49	35	279	2018-Q2	160	56	21	301
2018-Q3	114	6	4	153	2018-Q3	60	2	0	130
2018-Q4	189	81	29	277	2018-Q4	149	33	21	280
2019-Q1	264	59	48	248	2019-Q1	217	49	51	235
2019-Q2	359	46	65	167	2019-Q2	295	29	43	165
2019-Q3	180	0	14	66	2019-Q3	153	0	11	85
2019-Q4	296	10	35	123	2019-Q4	240	8	32	107
合計	2,599	510	389	2,788	合計	2,106	363	321	2,889

5 結論と今後の課題

5.1 結論

本研究の目的は、①金融機関が取引先の資金ニーズを事前に察知すること、および②資金ニーズを抱えている取引先へ主体的に訪問・営業を行うことについて、それぞれの方法論を示すことである。これらを達成することで、リレーションシップバンキング機能を有する中小・地域金融機関は、自機関が保有する財務諸表データから、資金ニーズを有する可能性が高い事業者を特定し、地域の資金ニーズに応じたリレーションシップバンキング業務を行うことが可能となる。

1 つ目の目的については、水戸信用金庫の秘匿化加工された取引先事業者の財務諸表データを用いて、長期借入金の増減を判別する数理モデルを構築することで達成された(4 節 1 項と 4 節 2 項を参照)。初めに、罇ほか(2022)で提示された方法論 2(表 3 参照)を水戸信用金庫の取引先事業者の財務諸表データに適用した。その場合、全体として ACC の期待値+16.2%ポイントの判別性能を示すモデルが構築された(データ分割前の ACC の期待値は 0.547)。水戸信用金庫へのヒアリングでは、「概ね 70%ほどの精度で長期借入金の増減を当てることができれば良いだろう」とのことであったため、方法論 2 は中小・地域金融機関が保有する財務諸表データに対しても、適用可能であることが示唆された。しかし、この結果は、罇ほか(2022)で確認された ACC の期待値からの判別性能の向上幅よりも小さく、70%の正解率を下回るケースも確認された。そこで、本研究では、方法論 2 に四半期でデータを分割する処理を加えた方法論 3(表 13 参照)を提案した。方法論 3 に則った処理を行うことで、中小企業の財務諸表データから、ACC の期待値+32.0%ポイントの判別性能を示すモデルが学習された。また、中小・地域金融機関の取引先事業者は、財務諸表に記載されていないマクロ経済状況や経営環境の変化に敏感である可能性が示唆された。ただし、本研究の実験のみでは、方法論 3 が方法論 2 より優れていると断言することはできないことに注意する必要がある。

なお、本研究で提案した方法論は、Imakura and Sakurai(2020)によって提案されたデータコラボレーション解析(Data Collaboration Analysis)について、シングルサイロ解析に相当するものである。データコラボレーション解析は、SVD 等の次元削減手法を用いて、機密性の高いデータを不可逆変換した中間表現データを外部の分析機関に送信し、分析機関は、1 つないしは複数の機関から収集したデータを解析することができる技術である。本研究は、中小・地域金融機関が保有する実データを用いて、データコラボレーション解析(Bogdanova et al. 2020; Takahashi et al. 2021; Imakura et al. 2021)を部分的に応用した嚆矢的な研究として位置づけることもできよう。

2 つ目の目的については、方法論 3 に則って構築したモデルが出力した結果を、営業活動に利用する際の例を示すことで達成された。具体的には、金融機関は、財務諸表データから長期借入金増減判別モデルを構築した後に、適切な閾値を設定し、モデルの出力を 0 もしくは 1 に変換する。次に、変換後の数値に基づく混同行列(図 2 参照)を作成する。そして、偽陽性に該当した事業者に対して、営業のリソースを集中する(図 4 参照)。ここで、偽陽性の事業者とは、モデルは長期借入金が増加すると判断したが、実際は増加しなかった事業者である。したがって、偽陽性の事業者は、他の事業者と比較した際に、より高い確率で資金ニ

ーズを有すると考えられる。ただし、偽陽性の事業者に金融機関のリソースを割くことによって、真に中小・地域金融機関の収益力と健全性の維持・向上がなされるのかについては、実証研究や実地調査による究明が求められる。

本研究の社会的貢献は、現実の中小・地域金融機関が保有する取引先事業者の財務諸表データから、当該事業者の長期借入金の増減を十分な精度で判別する方法論を提案し、資金ニーズを有する事業者へ接近するための具体例を示した点である。本研究で提示した手法を用いて、全国の中小・地域金融機関は、営業エリア内の資金ニーズを読み解き、リレーションシップバンキングとして、収益力と健全性の維持・向上に努めることが期待される。また、学術的な貢献は、財務諸表データから長期借入金の増減を判別するタスク設定において、罇ほか（2022）で提案された方法論2に、四半期によるデータ分割処理を付加した方法論3の適用によって、既存の判別性能を上回る結果を示した点である。加えて、データコラボレーション解析におけるシングルサイロ解析として容易に応用可能な方法で、リレーションシップバンキング機能を向上させるための方法論を提案した貢献も指摘しておきたい。

5.2 今後の課題

本研究で使用したデータは、水戸信用金庫の取引先事業者に限定されていることに注意する必要がある。つまり、いかなる中小・地域金融機関が保有する取引先事業者（主に中小企業と考えられる）の財務諸表データに対しても、方法論3の適用により、高い判別性能を示すモデルが学習されることは明らかではない。したがって、今後の研究では、多数の中小・地域金融機関が保有する取引先事業者の財務諸表データを集約し、長期借入金の増減を判別するモデルを構築することが求められる。しかし、金融機関が保有する取引先事業者の財務諸表データは機密情報に当たるため、複数の金融機関が保有する財務諸表データを金融機関の外部へ持ち出して、統合したうえで分析を行うことには高いコストを要する。また、このような統合分析を、金融機関の一職員が実施することは困難であろう。この点に関して、各金融機関で分散して保有されている財務諸表データに対してデータコラボレーション解析を応用することで、多くの中小企業のサンプルを用いた頑健な長期借入金増減判別モデルを費用対効果良く取得可能であると考えられる。

次に、本研究で作成した長期借入金に関する情報のさらなる利活用について検討する。例えば、モデルの出力を踏まえたうえで、各企業の財務分析を行うことができれば、より具体的な営業支援情報を考案することができるだろう。したがって、本研究で提案した方法論に基づくモデルを、ビジネスインテリジェンス（Business Intelligence、以下 BI）へ統合することが良い道筋であると思われる。BIとは、データの収集・解析・表示を行うことで、意思決定を支援するツールである（Rikhardsson and Yigitbasioglu 2018; Božič and Dimovski 2019; Ain et al. 2019）。また、実務においては、モデルの振る舞いを知ることによって、より効果的な営業支援情報を作成することができる可能性がある。したがって、本研究が示すもう一つの道筋は、モデルの出力結果の解釈を提供するための方法論を確立することである。具体的には、データコラボレーション解析に SHAP（Lundberg and Lee 2017; Lundberg et al. 2020）を応用した Bogdanova et al.（2022）の活用が挙げ

られる。ただし、ブラックボックスな機械学習モデルと本質的に解釈可能なモデルとの差を認識したうえで、分析を進めることが求められる (Rudin 2019)。

最後に、本研究で構築したモデルは、使用した学習データへの依存度が非常に高い。すなわち、ある分割データで学習したモデルでは、別の分割データの目的変数を精度良く判別することができない。したがって、いかなる財務諸表データに対しても、精度良く判別が可能なモデルを開発する手法が求められる。これに付随して、本研究は、厳密な機械学習の意味での未来予測を行っていない。未来予測を行うためには、少なくとも以下の修正が必要である。1つ目は、本研究で提案した方法論3に含まれる標準化やSVDによる次元削減処理を、学習データとテストデータに分割する以前に一括で適用している点である。この手順は、本研究で提案した判別モデルの利用例であれば問題ないが、テストデータが手元にない状況を想定すると不適切である。2つ目は、方法論として、目的変数との相関を用いた変数選択を使用した点である。この処理は、純粹な未来予測を目的とすると、将来得られるはずの情報をモデル構築のために使用しており、単純に $t+1$ 年度の予測能力を測定したい場合には、不適切であると考えられる。本研究のアプローチは、あくまでも、PDCA サイクルの Check のための AI を構想しているのである。したがって、より長期の予測を行い Plan に活かす場合は、これら2点を修正した方法論の提案が求められる。

参考文献

- 岡田幸彦・罇涼稀・秦涼太・今倉暁・櫻井鉄也. 2021. 「財務諸表は長期借入金の増加予測に資するのか?」『日本会計研究学会第80回全国大会自由論題報告』
- 金融審議会金融分科会第二部会. 2003. 『リレーションシップバンキングの機能強化に向けて』
- 金融庁. 2021. 『中小・地域金融機関向けの総合的な監督指針』
- 首藤昭信. 2019. 「AIが会計学研究に与える影響」『会計』195(2): 127-141.
- 滝川好夫. 2007. 『リレーションシップ・バンキングの経済分析』税務経理協会.
- 中小企業庁. 2021. 『2021年度版中小企業白書』
- 罇涼稀・秦涼太・今倉暁・櫻井鉄也・岡田幸彦. 2022. 「財務諸表データを用いた資金ニーズの見直しチェック AI の開発」『Department of Policy and Planning Sciences Discussion Paper Series』1383.
- Ain, N., G. Vaia, W. H. DeLone and M. Waheed. 2019. Two decades of research on business intelligence system adoption, utilization and success – A systematic literature review. *Decision Support Systems* 125: 113113.
- Altman, E. I. 1968. Financial Ratios: Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance* 23(4): 589-609.
- Berger, A. N. and G. F. Udell. 1995. Relationship Lending and Lines of Credit in Small Firm Finance. *The Journal of Business* 68(3): 351-381.
- Bogdanova, A., A. Imakura, T. Sakurai, T. Fujii, T. Sakamoto and H. Abe. 2022. Achieving Transparency in Distributed Machine Learning with Explainable Data Collaboration. *arXiv*: 2212.03373.

- Bogdanova, A., A. Nakai, Y. Okada, A. Imakura and T. Sakurai. 2020. Federated Learning System without Model Sharing through Integration of Dimensional Reduced Data Representations. *Proceedings of FL-IJCAI 2020*: 1-7.
- Boot, A. W. A. 2000. Relationship Banking: What Do We Know? *Journal of Financial Intermediation* 9 (1): 7-25.
- Božič, K., and V. Dimovski. 2019. Business intelligence and analytics for value creation: The role of absorptive capacity. *International Journal of Information Management* 46: 93-103.
- Byanjankar, A., J. Mezei and M. Heikkilä. 2021. Data-driven optimization of peer-to-peer lending portfolios based on the expected value framework. *Intelligent Systems in Accounting, Finance and Management* 28 (2): 119-129.
- Cessie, S. L. and J. C. V. Houwelingen. 1990. Ridge Estimators in Logistic Regression. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 41 (1): 191-201.
- Chittenden, F., G. Hall and P. Hutchinson. 1996. Small Firm Growth, Access to Capital Markets and Financial Structure: Review of Issues and an Empirical Investigation. *Small Business Economics* 8: 59-67.
- Elyasiani, E. and L. G. Goldberg. 2004. Relationship lending: a survey of the literature. *Journal of Economics and Business* 56: 315-330.
- Imakura, A., H. Inaba, Y. Okada and T. Sakurai. 2021. Interpretable Collaborative Data Analysis on Distributed Data. *Expert Systems with Applications* 177: 114891.
- Imakura, A. and T. Sakurai. 2020. Data Collaboration Analysis Framework Using Centralization of Individual Intermediate Representations for Distributed Data Sets. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* 6 (2): 04020018.
- Jakšič, M. and M. Marinč. 2017. Relationship Banking and Information Technology: The Role of Artificial Intelligence and FinTech. *Risk Management* 21: 1-18.
- Lobo, J. M., A. Jiménez-Valverde and R. Real. 2007. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography* 17 (2): 145-151.
- Lo'pez-Gracia, J. and F. Sogorb-Mira. 2008. Testing trade-off and pecking order theories financing SMEs. *Small Business Economics* 31: 117-136.
- Lundberg S. M., G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal and S. Lee. 2020. From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence* 2: 56–67.
- Lundberg S. M. and S. Lee. 2017. A unified approach to interpreting model predictions. *Proceedings of the 31st International Conference on Neural Information Processing Systems*: 4768–4777.
- Mandrekar, J. N. 2010. Receiver Operating Characteristic Curve in Diagnostic Test Assessment. *Journal of Thoracic Oncology* 5 (9): 1315-1316.
- Michaelas, N., F. Chittenden and P. Poutziouris. 1999. Financial Policy and Capital Structure Choice in U.K. SMEs: Empirical Evidence from Company Panel Data. *Small Business Economics* 12: 113–130.
- Myers, S. C. 1984. Capital Structure Puzzle. *NBER Working Paper*: w1393.

- Nosratabadi, S., A. Mosavi, P. Duan, P. Ghamisi, F. Filip, S. S. Band, U. Reuter, J. Gama and A. H. Gandomi. 2020. Data Science in Economics: Comprehensive Review of Advanced Machine Learning and Deep Learning Methods. *Mathematics* 8 (10): 1799.
- Öhman, P. and D. Yazdanfar. 2017. Short- and long-term debt determinants in Swedish SMEs. *Review of Accounting and Finance* 16 (1): 106-124.
- Okada Y. and N. Inamizu. 2014. Effect of Job Type on Perspective Index. *Annals of Business Administrative Science* 13 (6): 315-328.
- Ou, J. A. and S. H. Penman. 1989. Financial Statement Analysis and The Prediction of Stock Returns. *Journal of Accounting and Economics* 11: 295-329.
- Perboli, G. and E. Arabnezhad. 2021. A Machine Learning-based DSS for mid and long-term company crisis prediction. *Expert Systems with Applications* 174 (15): 114758.
- Rikhardsson, P. and O. Yigitbasioglu. 2018. Business intelligence & analytics in management accounting research: Status and future focus. *International Journal of Accounting Information Systems* 29: 37-58.
- Rudin, C. 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *nature machine intelligence* 1: 206–215.
- Stewart, G. W. 1993. On the Early History of the Singular Value Decomposition. *SIAM Review* 35 (4): 551-566.
- Takahashi Y., H. Chang, A. Nakai, R. Kagawa, H. Ando, A. Im akura, Y. Okada, H. Tsurushima, K. Suzuki and T. Sakurai. 2021. Decentralized Learning with Virtual Patients for Medical Diagnosis of Diabetes. *SN Computer Science* 2: 239.
- Youden, W. J. 1950. Index for rating diagnostic tests. *Cancer* 3 (1): 32-35.
- 一般社団法人全国信用金庫協会「信用金庫の制度」<https://www.shinkin.org/shinkin/seido/>（閲覧日 2021 年 10 月 19 日）
- NASA「Technology Readiness Level」https://www.nasa.gov/directorates/heo/scan/engineering/technology/technology_readiness_level（閲覧日 2021 年 10 月 19 日）
- NumPy API reference「numpy.linalg.svd」<https://numpy.org/doc/stable/reference/generated/numpy.linalg.sv-d.html>（閲覧日 2021 年 10 月 19 日）
- scikit-learn API reference「sklearn.linear_model.LogisticRegression」https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html（閲覧日 2021 年 10 月 19 日）

（付記）本研究は、NEDO 委託事業「データコラボレーション解析による生産性向上を目指した次世代人工知能技術の研究開発」、および JSPS 科研費（22H00895）による研究成果の一部である。

補足資料

本研究に後続する研究のための参考資料として、追加で実施した分析結果を報告する。

A. ①値モデル, ②比率モデル, ③差分モデルのそれぞれによる判別性能の比較

4 節 2 項のモデルは、財務諸表の値（現金・預金額や総資産額など）に関する説明変数と、比率（現金・預金額/総資産額など）に関する説明変数、および差分（ t 年度の総資産額 - $t-1$ 年度の総資産額など）に関する 3 種類の説明変数がモデルに組み込まれていた。しかし、比率は値から作成され、差分は値や比率から作成されるため、値のみで学習したモデルと、比率のみで学習したモデル、および差分のみで学習したモデルは、値・比率・差分のすべてを用いて学習したフルモデルと同等の判別性能を発揮する可能性がある。一方、財務諸表分析では、財務諸表の値に頼る分析には限界があり、「評価されるもの/評価するもの」のように、比率で分析する方が合理的である（新田ほか 2014, p.29-33）とされているため、比率は値よりも目的変数を良く判別する変量となることも考えられる。

そこで、値を示す説明変数のみで学習したモデル（以下、値モデル）と、比率を示す説明変数のみで学習したモデル（以下、比率モデル）、および差分を示す説明変数のみで学習したモデル（差分モデル）の判別性能を比較するアブレーション研究（Jørgensen and Igel 2021）を行った。説明変数の作成法を除き、データの前処理や適用する方法論は、4 節 2 項の実験と同様である。ただし、前処理後の値モデルの説明変数は 231 個であり、比率モデルの説明変数は 227 個であった。また、差分モデルの説明変数は 151 個であった。そのため、方法論 3 を適用するにあたり、SVD による次元削減後の次元数は、値モデルと比率モデルは 220 次元として、差分モデルは 150 次元とした。上記の設定のもとで行ったアブレーション研究の結果を表 A に示す。

表 A 値モデル, 比率モデル, 差分モデルのテストデータにおける ACC (加重平均値)

基準年度 t	値モデル		比率モデル		差分モデル	
	事業者数	ACC	事業者数	ACC	事業者数	ACC
2016	969	0.770	965	0.751	968	0.760
2017	3,585	0.789	3,578	0.780	3,582	0.766
2018	3,723	0.769	3,712	0.765	3,723	0.767
2019	3,702	0.777	3,695	0.769	3,700	0.762
加重平均値		0.778		0.769		0.765
フルモデルとの差異		-0.093		-0.102		-0.106
Exp ACCとの差異		+0.231		+0.222		+0.218

フルモデルは、値を示す説明変数と、比率を示す説明変数に加え、差分を示す説明変数のすべてを用いて学習したモデルのことである。

表 A について、Exp ACC は、データ分割前の目的変数の分布から計算したものである (Exp ACC = 0.547)。実験の結果、値モデル・比率モデル・差分モデルを問わず、全ての分割データにおいて、Train ACC と Test ACC は Exp ACC を上回り、Train AUC と Test AUC は 0.5 を上回った。これより、値のみ、比率のみ、あるいは差分のみを説明変数として学習に用いる場合であっても、期待値以上の判別性能を示すモデルが学習されることが示唆された。しかし、当初の想定とは異なり、判別性能に関して、値モデル・比率モデル・差分モデルのそれぞれが示した判別性能は、すべての変数を用いたフルモデルによる判別性能には及ばなかった (-10%ポイント程度の低下)。しかし、ACC の期待値からの精度向上幅は+20%ポイント以上であり、実務上要求される最低限度の判別性能は満たしていた。また、値モデル・比率モデル・差分モデルのそれぞれが示した判別性能に大きな差は見られなかった。この点について、値・比率・差分には、説明変数として、それぞれが目的変数を説明する固有の情報を有する可能性があることが示唆された。しかしながら、本研究は線形回帰モデルに適した変数選択を行っているため、非線形の関係を扱うことが可能なカーネルロジスティック回帰モデルや決定木系のモデルを用いて学習を行うことで、各モデルの判別性能に差が生じることも考えられる。ただし、本研究で提案した方法論に則る場合は、値・比率・差分のすべてを説明変数としてモデルの学習に組み込む方が、各々でモデルを構築するよりも、判別性能の向上が見込まれるという結果であった。

補足資料の参考文献

- 新田忠誓・善積康夫・辻峰男・木村晃久・中村亮介・木村将之. 2014. 『実践財務諸表分析：財務諸表の読み方と分析法』中央経済社.
- Jørgensen, R. K. and C. Igel. 2021. Machine learning for financial transaction classification across companies using character-level word embeddings of text fields. *Intelligent Systems in Accounting, Finance and Management* 28 (3): 159-172.