

論文

企業の実態把握のための新たな 分析手法の可能性

—不正会計の検知を中心とした文献のサーベイ—

上枝 正幸

キーワード

企業の実態
新たな分析手法
統計的手法
計算的手法
会計報告書
量的データ
質的データ
データマイニング
不正会計／会計不正
会計コミュニケーション

目次

- はじめに
- 企業の実態把握—特に、不正会計の検知—のための
新たな分析手法の先行研究の議論
(1) Ngai, Hu, Wong, Chen, and Sun (2011)
(2) Sharma and Panigrahi (2012)
(3) Gray and Debreceeny (2014)
(4) Abdallah, Maarof, and Zainal (2016)
(5) West and Bhattacharya (2016)
(6) Amani and Fadlalla (2017)
(7) Gepp, Linnenluecke, O'Neill, and Smith (2018)
- 不正会計の検知に関する最近の研究
- 結論および議論

注

【参考文献】

1. はじめに

本稿は、会計関連の報告書記載の情報を中心とする量的・質的データを用いて企業の実態を把握し、あるいは何らかの将来予測を導くことを目的とする「新たな」分析手法を検討する近年の文献を跡づける。前拙稿（上枝 2018）では、会計報告書、具体的にはアニュアル・レポート内の「経営者による財務・経営成績の分析（Management Discussion and Analysis, MD&A）」の言語（テキスト）部分の統計分析を通じて、不正会計¹の存否やその開始時期の特定・予測を企図する Purda and Skillicorn (2015) を主として取り上げ、将来研究の可能性を探った。本稿は、上枝 (2018) と学問上の関心の軌を一にしているが、(1) 言語（テキスト）のみではなくより幅の広い量的・質的データを分析の対象として含め、さらに (2) 分析の主たる目的、すなわち把握したいと望む企業実態に関しても、不正会計の検知を第一の論点にするものだけに限定するものではない²。企業の実態に迫るためには、分析の対象たるデータや目的を狭く絞りこむことよりも、それを可能な限りにおいて広く捉えることが少なくとも現時点では有用であると考えからである。それは当然のこと、上枝 (2018, 32 頁) において指摘したように、(3) 分析にあたって依拠するアプローチにとっても該当する。そうであるならば、さらに、上枝 (2018) における議論の大前提、すなわち、「将来の帰結をも含む企業の実態が会計関連の報告書における何らかの情報に写像・マッピングされるならば、当該会計報告書の読解と解析から企業の実態の把握が可能になる」という図 1—上枝 (2018, 30 頁) の図 2 の再掲—の仮定（モデル）をも超越し、「会計」の看板さえ取り払い、(4) 関心をもち知りたいと望む企業の実態に迫りえる情報であれば、その出処や属性は問わないものとするべきかもしれない。しかしながら、本稿はわれわれの会計コミュニケーション研究（上枝 2014a/b/c, 2016a/b）の一環、すなわち会計関連のディスクロージャー（情報開示）は企業実態の何をどの程度まで意思伝達できているのかを探求する試みの一つでもあることから、「会計関連の報告書記載の情報を中心とする」という文言については取り扱うことなく、本（第 1）節の冒頭に付けたところである。同様に、副題とした「不正会計の検知」については、会計コミュニケーションにおいて、不正会計の存在はコミュニケーション不全（miscommunication or communication failure）の最たるものであるといえ、その社会・経済的な悪影響が多種多様かつ金銭的に巨額に及び、よって相当数の興味深い先行研究が既になされてきており、したがって文献サーベイになじむ、あるいはすでに先行研究の暫定的な知見に基づくサーベイ論文が海外ジャーナルに登場していることによるものである。

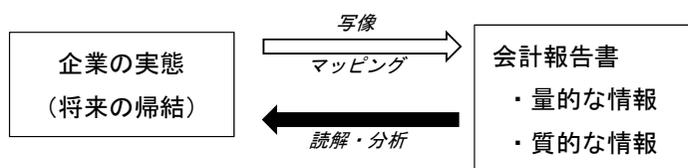


図 1. 企業の実態の表出と読解・分析・予測（上枝 2018, 図 2, 30 頁）

ところで、本稿でみる「新たな」分析においては、一般的に、いわゆるビッグデータ³や AI (artificial intelligence, 人工知能) などに係る技術を適用し、統計的・計算的な処理を通じ、データと現

実世界の事象—たとえば、企業の実態—との間にある何らかの法則性やパターンを見出そうとする。それは、次（第2）節でみる West and Bhattacharya (2016, 47-48) を参照して簡潔に言えば、インテリジェントないし計算知能的な (computational intelligence (CI)-based) テクニックであり、具体的には、統計的・計算的なデータマイニングの手法を利用するものである⁴。企業内部・外部の利害関係者の経済的意思決定にとって有用な情報の提供という会計の目的に変化はなくとも⁵、技術の進歩やパラダイムシフトにより、「新たな」アプローチが要求される、あるいは少なくとも「新たな」アプローチが旧来のアプローチよりも有効になるというのはごく自然な発想であると考えられる。Elliott (1992) は、たとえば、情報テクノロジー (information technology, IT) の進展が有史以来第三の波として経済社会に到来し、ビジネス、企業経営、延いては会計・監査実務を変え(て)ると論じた⁶。以来四半世紀余を経てなお技術の進展・環境の変化は継続中といえるが、経営や会計・監査のこれまでの変化に対応する企業の実態の読解・分析の「新たな」手法があるならばそれはいかなるものであろうか、というのが本稿の視点である。

折しも、わが国においても、ビッグデータや AI など分析技術の進展が会計・監査実務および会計学研究のさまざまな局面に及ぼしている、あるいは将来的に及ぼすと予想される広範な影響やその対応策に関する議論が盛んになってきている (大城 2014, 日本会計士協会 2016/2019, 宋 2018, 瀧ほか 2018, 日本経済新聞 2018 年 10 月 8 日付朝刊, 首藤 2019, 矢口 2019 など)。まさにかようなホットトピックにつき、本稿は、海外文献を中心として関連する学術研究をサーベイし、今後の指針を探ろうとするものである。

本稿の位置付けは以上の通りであるが、本稿の構成は次のようである。本（第1）節に続く第2節は、本稿の主要部分であり、企業の実態把握のための新たな分析手法に関し、先行研究でなされている議論をレビューする。第3節は、前（第2）節では取り上げない「個別の」研究を取り上げて簡潔にみる。なお、前拙稿 (上枝 2018, 第2節, 18-20 頁) でも本稿のテーマと関連する文献をすでに扱っていることから、これら二つの節（第2節と第3節）ではなるべく前拙稿との重複を避けるように心がける。最終の第4節では、議論をまとめ結論を述べる。

2. 企業の実態把握—特に、不正会計の検知—のための新たな分析手法の先行研究の議論

本（第2）節は、企業の実態把握のための「新たな」データ分析手法の先行研究の議論をみる。そのさい、本（第2）節では、個別の論文—具体的には、Purda and Skillicorn (2015)—を取り上げて吟味した上枝 (2018) のようなアプローチとは相違し、あるいはそれは次（第3）節にゆずり、統計的・計算的なデータマイニングの手法を会計・監査の問題への適用に関する論点自体を包括的・全体的に議論する先行研究を扱うものとする。具体的には、公刊の年の順番に、(1) Ngai et al. (2011)、(2) Sharma and Panigrahi (2012)、(3) Gray and Debreceeny (2014)、(4) Abdallah et al. (2016)、(5) West and Bhattacharya (2016)、(6) Amani and Fadlalla (2017)、および (7) Gepp et al. (2018) の七つの文献を幾分恣意的に選択した⁷。なお、当該領域の先行研究の特徴を簡潔に示せば、

これら7論文のうち、監査に対するデータマイニングの適用を検討し、執筆当時までの研究の分類学 (taxonomy) を示した Gray and Debreceeny (2014, 357) では、発展の初期段階かついわば「手あたり次第の (scatter-shot)」アプローチがとられているところであると表現されている。2019年現在においても、筆者にとって、状況に大きな変化はないように感じられることから、7論文の議論をさらに要約するというよりむしろ、公刊年の順にみることにした。また、金融不正 (financial fraud) 全般の検知という幅広い論点を扱う、たとえば Ngai et al. (2011) などの研究も含まれるものの、前(第1)節において述べたように、社会的・経済的な悪影響の大きさおよび会計コミュニケーション—あるいはコミュニケーション不全—に対する筆者の個人的な関心に鑑み、「会計不正 (accounting fraud) ないし不正会計の検知」の論点が分析の主たる目的となる。

(1) Ngai, Hu, Wong, Chen, and Sun (2011)

本論文は、金融不正の検知を目的としてデータマイニングの技法を適用する先行文献を包括的にレビューし、関連研究の分類のための適切なフレームワークを構築し、もって将来研究の道標を導こうとする、彼らによれば従来には存在しなかった初めての試みである (Ngai et al. 2011, 559-560)。彼らはまず、1997年から2008年までの12年間に公刊された論文を対象として、七つのオンライン・データベースからキーワード—金融不正、データマイニングおよびビジネスインテリジェンス—検索し、ヒットした約1,200本のなかから49本の関連論文を絞り込んで考察する。結果的に、金融不正は、四つのカテゴリー、すなわち①銀行不正、②保険不正、③証券と商品の不正、および④関連するその他金融不正に分類され、会計不正—財務情報の偽造 (falsification of financial information) や不正な財務諸表 (fraud financial statements, FFS) と表記—はこのうち④関連するその他金融不正に含まれる企業不正 (corporate fraud)⁸ の一つとされる (Ibid., 561-562)。また、金融不正に適用されるデータマイニング技法としては、①分類 (classification)、②クラスタリング (clustering)、③予測 (prediction)、④外れ値検知 (outlier detection)、⑤回帰 (regression)、および⑥ビジュアル化 (visualization) の六つがあるとする (Ibid., 562)。なお、具体的には、26個のデータマイニングの技法が掲げられ、なかでもロジスティック・モデル (75の適用事例のうち、21.3%、以下全て)、ニューラルネットワーク (13.3%)、ベイジアン信念ネットワークと意思決定木 (ともに、6.7%) の順で頻繁に利用されてきているとする (Ibid., 563-564, and Table 7, 566)。彼らの分類のフレームワークは、金融不正の論点 (①から④) およびデータマイニング技法 (①から⑥) を縦横の軸とするものであり、したがって、先の49の先行研究は計24個 (=④×⑥) のセグメントのうちのどこかに収容される (Ibid., Fig. 3, 563)。彼らによると、会計不正を取り扱った論文数は6本、うち五つが分類、残る一つが回帰のデータマイニング技法を用いているとする (Ibid., Tables 5-7, 565-566)。

以上、Ngai et al. (2011) からは、金融不正の論点 (①から④) とデータマイニングの技法 (①から⑥) の二つの軸による先行研究の分類のためのフレームワークにより、どの研究領域が既に実施され、あるいは相対的に未開拓なままであるかが明確に理解できる。さらに、彼らの結論部分 (Ibid., 565-567) では、当該分野の研究上の課題が提示される。たとえばわれわれの関心に沿うところでは、

(i) 不正な財務諸表、すなわち不正会計を検知する研究にとって必要十分なデータを入手することの困難さ (Fanning and Cogger 1998) であり、(ii) 研究者と実務家との間の視点のギャップ—すなわち、研究者はデータマイニングの技法やモデル自体に関心があるのに対し、実務家は実務上の原則や問題解決法を知りたいと望むという差異—であり、さらに (iii) 分析のコスト感応度 (cost sensitivity) —不正検知のさいの偽陰性や偽陽性の誤謬の発生によるコスト—を研究者らが考慮する必要性などである。

(2) Sharma and Panigrahi (2012)

本論文は、財務会計不正の検知 (financial accounting fraud detection, FFAD) のためのデータマイニングの技法の適用に関する包括的なレビューを実施し、そのフレームワークを提示することを目的とする (Sharma and Panigrahi 2012, 37)。彼らはまず、不正 (fraud) や金融不正 (financial fraud)、また会計不正 (accounting fraud) を定義付けし、前小節 2 (1) でみた Ngai et al. (2011) による研究の分類を援用し、さらにデータマイニングの技法を記述する (Ibid., 38-41)。彼らの結論部分 (Ibid., 41) は、Ngai et al. (2011, 566-567) と酷似しており、よって追加的に参照すべき箇所は殆どないものの、財務会計不正の検知 (FFAD) という本稿の関心に合致する 2000 年代までの一連の研究の文献目録 (Ibid., Tables 1-4, 42-44) が適用技法ごとにまとめられ、さらに FFAD のためのデータマイニングの技法のフレームワークが図示 (Ibid., Fig. 2, 44) されており、こちらは関連研究の概要を把握するさいに有用である。後者のデータマイニング (DMと略記) の技法のフレームワーク (Fig. 2) については、以下において本稿の図 2 として再掲示している。

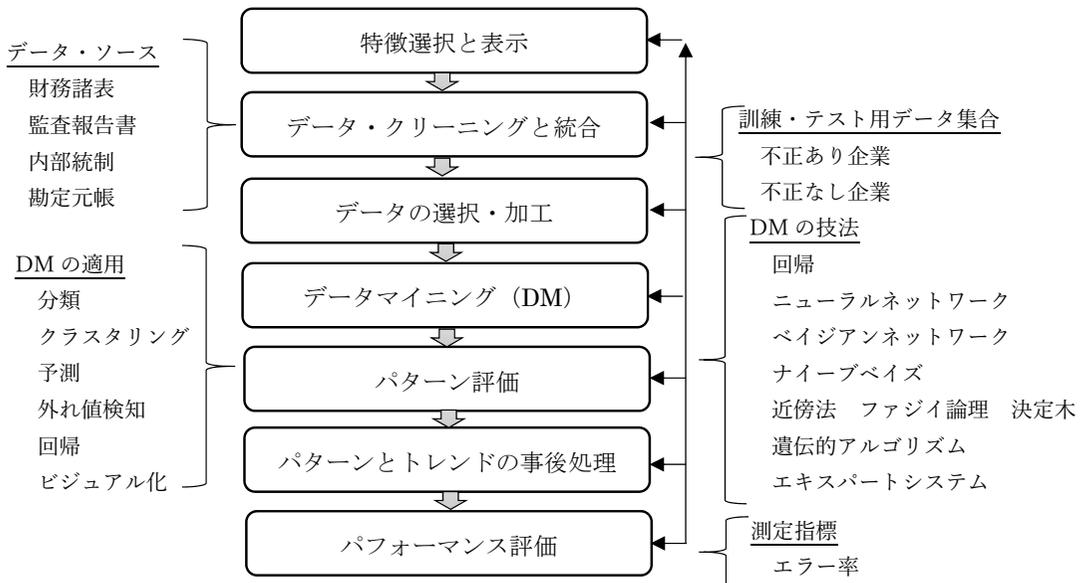


図 2. 財務会計不正のデータマイニングに基づくフレームワーク

(出典：Sharma and Panigrahi 2012, Fig. 2, 44)

(3) Gray and Debreceeny (2014)

Gray and Debreceeny (2014) は、財務諸表監査における不正検知に対するデータマイニングの適用の研究のための指針となるような分類学 (taxonomy) を提示しようとする。彼らは最初に、監査の七つの局面 (Ibid., Fig. 2, 364) と各局面においてデータマイニングが果たしうる潜在的な役割について議論し、監査手続上のデータ調査プロセス (Ibid., Fig. 3, 367) を示している。その後、Gao and Srivastiva (2011)⁹ を参照して、会計不正が行われるさいに経営者が企図するスキームにつき、(i) 勘定スキーム (account schemes) ないし不正のタイプ、すなわち「結果的に不正な財務諸表をもたらす、架空収益のような、勘定残高や開示を操作することにより、経営者が不正にコミットするスキーム」(Ibid., 1)、および (ii) 証拠スキーム (evidence schemes)、すなわち「監査人を欺き、よって不正を隠蔽するため、架空・偽造文書を用いることのように、経営者が証拠を創作する (あるいは隠蔽する) スキーム」(Ibid., 1) の二つの構成要素に分け、それぞれの具体的な内容の 63 (7 × 9) の組み合わせ—たとえば、(i) 架空収益計上のために、(ii) 文書を偽造するという組み合わせ、または (i) 架空資産を計上するために、(ii) 第三者と共謀するなど—による分類がなされる。最終的に彼らは、不正会計の勘定・証拠の両スキームの組み合わせが実際に生起する頻度とともに、それぞれの組み合わせに対するデータマイニングの適用可能性の三つの段階—低位、中位ないし高位—を一つの図 (Ibid., Fig. 7, 376) へと要約し、その分類学を終える。彼らの図 (Fig. 7) によれば、どのような不正会計のスキームが経営者にとられ、あるいはとられやすく、さらにそれぞれの不正会計のスキームに対するデータマイニングの適用可能性の程度が簡明に理解できるものとなっている。

彼らの論文では、さらに、上枝 (2018) でみたようなテキスト様式のマテリアル—ここでは、Eメールなど—や非財務情報のデータマイニングの適用例の簡潔なレビューや将来研究の可能性が示される (Gray and Debreceeny 2014, 363 and 377-378) ほか、監査現場における適用時の障害 (Ibid., 377-378) などの議論がなされている。

(4) Abdallah, Maarof, and Zainal (2016)

Abdallah et al. (2016) は、コンピュータ技術の進展と企業の成長により、電子商取引システムを介した金融取引がよりいっそう一般的になってきたことに伴い、合法的な利用者と不正行為者 (fraudsters) との同一市場内の併存が問題となってきているとし、両者を識別するための不正検知システム (fraud detection system, FDS) に関する先行研究をサーベイしている。彼らは、不正の定義付けをしたあと、本稿の 2 (1) 節でみた Ngai et al. (2011) や Richhariya et al. (2012)¹⁰ などを金融不正の検知のためのデータマイニングの技法の適用に関して先行する議論として紹介するほか、クレジット・カードの不正利用や医療・自動車保険の保険金の不正搾取をも含む幅広い論点の先行研究をレビューする。1994 年から 2014 年までの先行研究の具体的な検討をする彼らの第 6 節では、しかしながら、クレジット・カード不正、テレコミュニケーション不正、健康保険不正、自動車保険不正およびオンライン・オークション不正の五つが取り上げられており、本稿の関心事たる会計不正の問題は扱われていない。検知対象となる不正のタイプにより研究上の取り扱いが異なるとされる

(Abdallah et al. 2016, Fig. 11, 108) ことから、不正会計の検知に特化した場合、どのようなシステムを用い、(i) コンセプトドリフト (concept drift)、(ii) 分布の歪み (skewed distribution)、(iii) 巨大なデータ量 (large amount of data) (の縮減)、(iv) リアルタイム検知のサポート (supports real time detection) (Ibid., 94-97) という彼らが指摘する四つの課題¹¹にどのように取り組むかにつき、われわれが新たに検討する必要がある。これら課題への対応方法につき、本稿において俄かには論じえないものの、各個別の適用事例と不正会計の間にある類似する特徴の抽出が議論のスタート地点になるかもしれない。

(5) West and Bhattacharya (2016)

West and Bhattacharya (2016) は、前 (第1) 節における本稿の検討対象の明確化にさいして触れたように、インテリジェントないし計算知能的な (computational intelligence (CI)-based) テクニック、具体的には、統計的・計算的なデータマイニングの手法を利用する金融不正の検知に関するおよそ 50 の先行研究の分析結果を包括的にレビューしようとするものである。ここでも、会計不正—正確には、財務諸表不正 (financial statement fraud)—は、数多い金融不正のなかの一つとしての取り扱いとなる (Ibid., Sec. 3.3.¹², 51, and 59)。しかしながら、扱われる 12 個のデータマイニング技法の概要と特徴、およびそのうち財務諸表不正に適用されうる技法が明示されている (Ibid., Fig. 3, 51 and 51-56)¹³ ことから、当該分野の研究を概観するさいに非常に有用であるといえる。論文の第5節 (Ibid., 56-59) において彼らは、他の先行研究と同様の検知アルゴリズムと不正タイプの二つの基準に加え、Bhattacharyya et al. (2011, 608)¹⁴ を参照し、データマイニング技法間の不正検知のパフォーマンスの優劣を評価すべく、(i) 精度 (accuracy)、すなわち「全サンプル」のうちある技法により正確に分類されたサンプルと誤分類されたサンプルとの比率、(ii) 感応度 (sensitivity)、すなわち「不正ありのサンプル」のうち不正ありと正しく分類されたものと不正なしと誤って分類されたものの比率—真陽性・偽陽性比率とも言及—、および (iii) 特異性 (specificity)、すなわち「不正なしのサンプル」を正しく不正なしと分類したものに対する誤って不正ありと分類したものとの比率—いわゆる真陰性・偽陰性比率—という三つの規準を用いる分類方法を議論している。彼らの図9 (Ibid., Fig. 9, 60) (の一部) は、財務諸表不正の検知手法間のパフォーマンス比較を示しており、限られた研究数が存在するのみという現段階においては参考に止めるべきであろうが、インテリジェントないし計算知能的な技法の伝統的な統計手法に対する優位性が確認できるほか、当然のこと、これら手法間におけるパフォーマンスの差、よって優劣が存在する可能性が示唆される。

なお、当該研究領域の将来展望に関しては、金融不正が取り扱いに慎重を要するセンシティブなトピックであることからくるプライバシー保持の問題、不正する側の不正行為テクニックの進化のペースに検知手法の進歩が追いつかない場合の問題、不正のタイプを問わずに一般的・包括的に利用可能な検知手法を開発することの意義の大きさなどが議論されている (Ibid., Sec. 6, 60-63)。

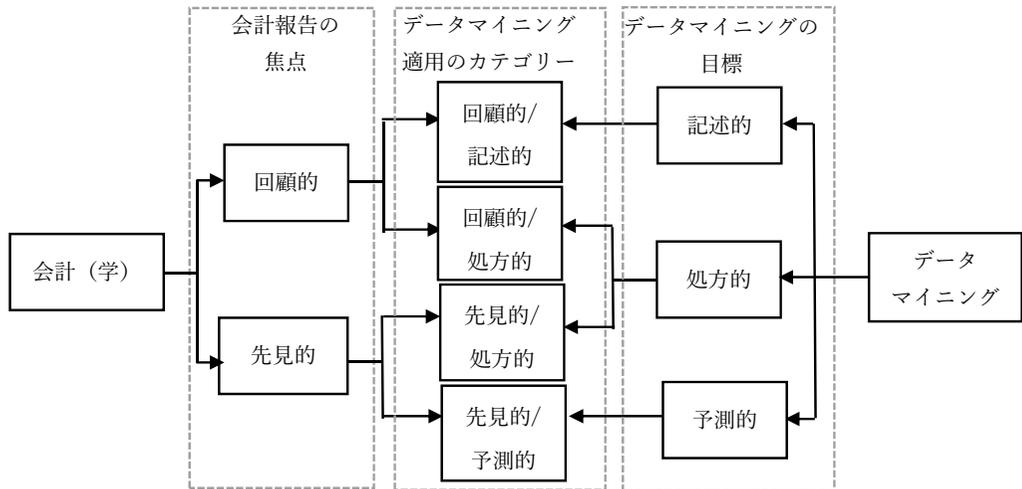


図3. 会計(学)におけるデータマイニング適用のフレームワーク

(出典: Amani and Fadlalla 2017, Fig. 2, 36)

(6) Amani and Fadlalla (2017)

Amani and Fadlalla (2017) は、「会計におけるデータマイニングの適用—文献レビューおよびフレームワーク系統化—」というその表題の通り、1989年以降2014年に至るまでの209という大量の先行研究について、特定のデータマイニングのパラダイムの利用に関してだけでなく、全体論的な(holistic)観点からレビューし、会計のさまざまな局面におけるデータマイニングの適用に関し、その執筆時点までの状況を明らかにし、さらに将来的な研究機会を考察するものである。彼女たちの集計によれば、(i) 同分野の文献は1995年から2001年にかけてと2004年から2014年にかけて—ピークは2013年—の二つの時期に増加傾向があること (Ibid., Fig. 3, 37)、(ii) データマイニング適用の目標としては予測的なもの(82%)が圧倒的に多く、記述的なもの(11%)と処方的なもの(7%)が続くこと (Ibid., Fig. 4, 38)、(iii) データマイニングのタスクとしては分類(140, 67%)が最大であり、以下推測(25, 12%)、クラスタリング(12, 6%)、さらに最適化・予測(それぞれ11・10、ともに5%)が続くこと (Ibid., Fig. 5, 38)、(iv) データマイニングの具体的な技法としてはニューラルネットワーク(適用数99)、回帰(適用数41)、決定木(適用数30)、サポート・ベクトル・マシン(適用数23)、遺伝的アルゴリズム(適用数22)、さらにテキストマイニング¹⁵(適用数15)と続くこと (Ibid., Table 1, 39)、さらに(v) 適用された論点別では保証とコンプライアンスが64%、管理会計が25%、また財務会計と会計情報システム(AIS)が11%であるとされる (Ibid., Figs. 6 and 7, 39 and 40)。彼女たちの提示するフレームワークは、**図3** (Ibid., Fig. 3, 36) のように示され、図の中央の列にあるデータマイニング適用のカテゴリー内の「回顧的(retrospective)/記述的(descriptive)」、「回顧的/処方的(prescriptive)」、「先見的(prospective)/処方的」と「先見的/予測的(predictive)」の四つの組み合わせの何れにある先行研究が該当するかを分類・集計し、またデータマイニングの手法を適用するさいに有用な会計(学)の論点を議論するのに用いられ、実際の適用(研究)例をマッピングした図9 (Ibid., Fig. 9, 50) が最終的に提示されている。

本稿でみてきている不正会計検知の論点に関していえば、保証とコンプライアンスに含まれる三つのサブカテゴリーの一つの「フォレンジックアカウンティング (forensic accounting)」—他の二つは、監査と企業の財務健全性—の小節 (Ibid., Sec. 5.3.4.3, 45-46) において先行研究が扱われており、適用の主目的としては「予測」、主要タスクは「分類」、適用例の多いカテゴリーとしては「先見的/予測的」、よく用いられる「技法」はニューラルネットワークと回帰であるとされる (Ibid., 46)。さらに、将来研究の機会が存する領域が列挙され、「経営者の特性に関する諸変数のインプットの拡充、分類器の統合に対するさまざまなアプローチの検証、さまざまな学習のアルゴリズムとモデル・アーキテクチャの検証、さまざまな時間粒度とデータ処理アプローチの探求、複数の企業・事業タイプに対するモデル展開の範囲の拡張、インプット変数の慎重な選択、予測のホライズンの延長、財務以外の変数の包含とビジュアル分析の増加、モデル比較に対する注意の傾注、金融不正予測時のデータマイニングとテキストマイニングとの結合など」 (Ibid., 46) が提示されている。さらに、CRISP-DM¹⁶ のフレームワーク内における会計 (学) のデータマイニングの適用の全般的な強み、弱みおよび推奨事項の三つが彼女たちの表2 (Ibid., Table 2, 47) に要約されており、当該領域の研究者にとって有用であるといえる。このほか、データマイニングの適用の領域の一つであるテキストマイニングは、会計報告書のテキスト部分の情報の豊潤さゆえに分析を実施する便益は大きいものの、それ自体が新しくかつ馴染みがないことにより、相対的に非常に少ない数の適用に留まっているという指摘がなされる (Ibid., 48)。最後に、われわれの関心と合致する記述としては、ニューラルネットワークといういわばブラックボックス的な一すなわち、説明力・説明可能性が相対的に低いとされる—モデルを導出する技法が一般的であることからくる研究上の限界、また財務諸表の虚偽記載が意図的なものか単なる誤謬なのかを区別するモデルの開発が有用でありかつデータマイニングにとって将来的に有望な研究領域と考えられるということの二つの指摘 (Ibid., 52)、さらにここで扱った保証とコンプライアンスのトピック以外に、財務会計の幅広い論点¹⁷ に対してデータマイニングが適用可能であり、かつ適用されてきたという事実 (Ibid., Fig. 7, 40, and 40-41) そのものがあると考えられる。

(7) Gepp, Linnenluecke, O'Neill, and Smith (2018)

Gepp et al. (2018) は、監査におけるビッグデータの技法 (big data techniques) の利用について包括的に考察するレビュー論文であるが、冒頭において彼らは、「他の関連諸領域におけるほどは実践は普及していない」 (Ibid., 102) という結論を述べている¹⁸。Gepp et al. (2018) 稿は、その第2節 (Ibid., Sec. 2, 103-104) においてビッグデータとビッグデータの技法を定義付け、補遺 (Ibid., Appendix A, 111-113) においては主たる技法に関する簡単な説明を供することにより、議論を開始する。続く第3節 (Ibid., Sec. 3, 104-108) は彼らの論文の中心部分であり、関連文献のレビューを体系的なものとするべく、300にも及ぼうかという大量の先行研究が探索され、監査における三つないし四つの研究の道筋、すなわち (i) 財務的な困窮 (度) のモデリング (financial distress modeling)、(ii) 金融不正のモデリング (financial fraud modeling)、(iii) 株式市場の予測と定量的モデリング (stock market prediction and quantitative modelling) の三つ、および (iv) 監査実務 (auditing) を加え

た四つが示される。本稿でみてきている不正会計の検知の論点はこのうち、(ii) 金融不正のモデリング (Ibid., Sec. 3.2, 105-106) において取り扱われ、本稿の 2 (3) 節でみた Gray and Debreceny (2014) や Van Den Bogaerd and Aerts (2011)、さらに上枝 (2018) でもみた Purda and Skillicorn (2015) の関心事であった会計報告書内のテキスト部分の分析も含め、先行研究が要約されている。当該分野に関し、彼らは、研究により分析からの結論が異なりうる初期段階であること、金融不正の発生の有無の二者択一のモデル化であるがゆえに、個別の事例ごとの「金融不正のコスト」が看過されている¹⁹ こと、さらに複数の不正行為者間の共謀がしばしば起こる事実が未考慮であることなどを、将来研究の機会であるとしている。自然言語処理 (natural language processing, NLP) はさらに、株式市況の動向の予測にオンライン・ニュースの利用が役立つなど、これまでにはなかった新規の情報源が付加価値をもたらす可能性があるがゆえに、監査分野の将来研究にとって便益が大きいと議論する (Ibid., 108-109)。

上記のほか、個人的に注目したい指摘としては、(i) 財務的な困窮 (度) のモデリング研究は、監査実務においては継続企業の前提 (ゴーイング・コンサーン) に関する意見の論点と関連するが、中小監査法人がクライアント喪失の恐れからその指摘を躊躇する、あるいは現実に行っているとしたら、ビッグデータの技法から導かれる結果は客観的な証拠として機能し—正確には、機能するならば—、いわゆる監査人の独立性の強化につながるというものがある (Ibid., 105)。こうした指摘自体さまざまな観点²⁰ からみて興味深いものの、不正会計の検知にも該当する—すなわち、会計不正ありという分析結果をもってしても、当該企業が認めない、あるいは詳細な調査を拒まれる場合に、ある程度は客観的な証拠となりうる可能性がある—ものであると考えられる。最後に、これまでの研究から導かれる重要な知見と将来的な機会として、単一の技法の利用よりも複数の技法を組み合わせることがずっとより有効であること、またビッグデータの技法は、熟練者、すなわち能力や経験のある監査人を代替するものではなく、その職務を補完するものである²¹ こと (Ibid., 108) の二つも、特に真新しいものではないとはいえ重要な指摘といえるであろう。

以上、本 (第 2) 節では、個別の論文のレビューは次 (第 3) 節へとゆずり、統計的・計算的なデータマイニングの手法の会計・監査の問題に対する適用に関する論点自体を包括的・全体的に論じた先行研究七つをみてきた。研究は依然としてその初期段階にあるものとはいえ、確たる知見が得られているわけでない。しかしながら、本 (第 2) 節のサーベイからは、当該分野の研究がこれまで扱ってきた論点、適用されてきた分析の技法など研究を分類するさまざまな観点やアプローチ、研究結果や分析の技法自体を測定・評価する基準、さらにこれまでに判明している研究の限界や将来展開の可能性など、今後研究を進めるうえで非常に有用な知識が得られたと考える。

3. 不正会計の検知に関する最近の研究

本 (第 3) 節は、前 (第 2) 節における包括的・全体的な先行研究の議論の概観に続き、会計関連の報告書からの情報を中心とする量的・質的データの新たな分析手法を用い、企業の実態、特に本稿では会計不正の有無に迫ろうとする「個別の」研究を簡単にみていくこととする。前 (第 2) 節の 7

文献は、さまざまな視角からの議論であること以外に、これまでなされてきた先行研究の優れた文献表ともなっている。それら文献表には、筆者（上枝）が通常はアクセスしないような、技術系の論文を採録する学術誌や近年になって創刊された新興の会計学の雑誌も含まれる。たとえば、不正会計検知のために財務データを主として用いるデータマイニングの研究には、ギリシャ企業を対象とする Kirkos et al. (2007) や中国・台湾企業を対象とする Ravisankar et al. (2011)、Huang (2013) および Chen et al. (2016) さらに二つのデータマイニング技法のパフォーマンスを比較する Tangod and Kulkarni (2015) がある。会計関連の報告書のテキスト部分についてもこれを分析のインプットとして用いる研究としては、米上場企業の年次報告書の「経営者による財務・経営成績の分析 (MD&A)」内のテキストを対象とする Glancy and Yadav (2011)、同じく米上場の中国企業の年次報告書の MD&A のテキストが対象の Dong et al. (2014)、さらに台湾企業の年次報告書のナラティブ開示を扱う Chen et al. (2017) があり、Hajek and Henriques (2017) は年次報告書内の財務情報と経営者コメント（テキスト）の双方から特徴となる変数を抽出する。前（第2）節でみたように、個別の論文の知見が当該論点に対して何らかの結論を与える段階ではないものの、われわれにとって興味深い記述が散見される。たとえば、Glancy and Yadav (2011) は、欺瞞検知 (deception detection) と題する小節 (Ibid., Sec. 2.2., 596-597) を設けて欺瞞の理論的な背景を考察し、「虚偽 (deceit) は、書き手が虚偽であることを知っているがゆえに探知可能である」(Ibid., 598) とする概念化を通して不正検知の計算モデルを構築する。また、不正の検知・特定に比べて、その重要性にもかかわらず不正の事前の抑止に係る研究が少ないこと (Tangod and Kulkarni 2015) や海外進出した中国企業の不正会計が大きな問題になっていた—よって、不正会計検知に関する研究が勃興した—という事実 (Dong et al. 2014) が指摘されている。このほか、Segal (2016) は、会計不正の検知と抑止のための先進的なテクノロジーの利用を論じるなかで、会計環境の主たる利害関係者に対するアンケート調査を実施し、(i) 不正リスク低減のための成功要因、および (ii) 会計不正抑止に向けた主たる課題を問う (Ibid., Sec. 3, 51-53) など、幅広い論点を考察する。

本（第3）節の最後に、Humpherys et al. (2011) をみておきたい。なぜなら同稿は、金融セクターのビッグデータ分析のテキストマイニングに関する最近の Bach et al. (2019) の文献レビューにおいて、もっとも引用数の多い論文の一つ—73回の被引用数であり第2位—とされる (Ibid., Table 1, 7 of 27)²² ためである。彼らは、負の影響が甚大であった2000年代の会計不正の発生、またヒトは一般的に不正検知が得意ではないという心理学の知見 (Bond and Depaulo 2006) などを受け、監査人や執行官 (enforcement officers) などの職務を扶助する「新たな」手法として財務諸表に対するテキストマイニングの適用を考察する (Ibid., 585 and 593)。意図的な隠蔽や共謀なども伴うがゆえに不正発見は難易度が高いことから、先見的な (proactive) アプローチが必要であるともされる (Ibid., 586)。彼らが着目したのは、利害関係者にもっとも読まれるものの、「準」強制的な開示でありかつ未監査である (Ibid., 585) 年次報告書内の「経営者による財務・経営成績の分析 (MD&A)」のテキスト部分である。彼らは、コミュニケーションと心理学の文献から不正と欺瞞に関する八つの理論と検知手法—1) 基準にもとづく内容分析 (Criteria-Based Content Analysis²³, CBCA)、2) 科学的内

容分析 (Scientific Content Analysis, SCAN)、3) リアリティ・モニタリング (Reality Monitoring, RM)、4) 経営者による不明瞭化仮説 (Management Obfuscation Hypothesis, MOH)、5) 情報操作理論 (Information Management Theory, IMT)、6) 対人的欺瞞理論 (Interpersonal Deception Theory, IDT)、7) 四因子論 (Four Factor Theory, FFT)、および 8) 漏洩理論 (Leakage Theory, LT) —を参照して (Ibid., 586-587)、虚偽の開示がなされる場合の八つの言語的特徴を仮説として導出する (Ibid., 586-589)。たとえば、虚偽情報を含む MD&A の書き手ほど、(1) 企業の実態よりも好ましい印象を読み手に誤ってもたせるよう、当たりの良い (pleasant) 語句、比喩 (imagery)、情緒 (affect) また活性化 (activation) に係る言語表現を用いる傾向が強くなり、(2) 事象や行動に対する経営者の責任を曖昧にするよう、受動態や不確実性を表わす言葉を多用し、(3) 説得力を高め、信憑性があるような外観を醸成するため、文章や動詞・修飾語の数をそれぞれ多くし、かつ用いる言語の特殊性 (specificity) が高いと考えるのである。彼らは次に、「会計及び監査に関する執行通牒 (Accounting and Auditing Enforcement Release, AAER と略記)」の公報²⁴ から不正会計のある様式 10-K の報告書 101 を抽出し、101 の不正会計のない—正確には、不正会計ありとして摘発されていない—報告書と対 (ペア) にした 202 の報告書サンプルを構築し、上記の言語的な特徴—具体的には、24 の変数を選択—が不正の有無を識別する手がかりとなりうるか、さらにはロジスティック回帰やナイーブベイズ、サポートベクトルマシンのような機械学習のアルゴリズムを含む五つの分類器のパフォーマンスを比較する。検証の結果、彼らは、不正なディスクロージャーは、そうでないものと比べ、活性化の言語、単語数、比喩、当たりの良い表現がそれぞれ多く用いられる一方、語彙の多様性 (lexical diversity) は低いこと、さらにナイーブベイズと C4.5 のアルゴリズムによる縮減モデルの分類精度が高いことを知見する。

以上、本 (第 3) 節では、会計関連の報告書の量的・質的データに新たな分析手法を適用することにより企業の実態、特に本稿では会計不正の有無を判定しようと企図してなされてきた一連の研究を簡単にみた。レビューは、前 (第 2) 節でみた文献から知り得た先行研究を中心としてなされ、論点に関する決定的な証拠が存在しない段階であるがゆえに、最後にみた Humpherys et al. (2011) を除くほかは、扱われている論点や将来研究の実施に際して配慮すべき指摘を中心として簡潔に行っている。なお、本稿の第 1 章の最終段落で述べたように、上枝 (2018, 第 2 節, 18-20 頁) の関連文献の紹介との重複を避けるように配慮している。

4. 結論および議論

本稿では、企業の実態に迫ろうと会計関連の報告書の量的・質的データを用いる、「新たな」分析手法を検討する近年の文献をみてきた。企業の実態としては、不正会計の検知を主目的としたものの、データの出所や属性および依拠するアプローチなどには特段の制約を設けることなくできる限り幅広く対象とするように心がけた。本稿でみた「新たな」分析手法に依れば、ビッグデータや AI (人工知能) などと関連する技術を用い、統計的・計算的な処理を通じ、量的・質的データと現世の事象との間の何らかの法則性やパターンを見出し、延いては将来予測や問題解決につながることを期待され

る。かような目的の達成のため、本稿は、第1節の導入部分に続き、統計的・計算的なデータマイニングの手法の会計・監査の問題に対する適用に関し、第2節は包括的・全体的に議論する七つの先行研究を、第3節は「個別の」研究をそれぞれ取り上げて簡潔にレビューした。これら二つの節の内容をまとめると、以下の3点になるだろう。第一に、「新たな」分析手法の適用に関する確たる結論は未提示であり、当該領域の研究はまだその初期の段階にあるという記載が散見された。第二に、先行研究は、(1) 会計不正の検知や倒産の予測などの論点、すなわち分析を適用する目標、および(2) データマイニング技法という二つの属性 (Ngai et al. 2011)、または(3) 回顧的・先見的という二つの会計報告の焦点、および(4) 記述的・処方的・予測的という三つのデータマイニングの目標の組み合わせ (Amani and Fadlalla 2017) などの基準により分類されうる。第三に、ある分析手法のパフォーマンスの優劣の評価は、(i) 精度、(ii) 感応度、および(iii) 特異性 (たとえば、Bhattacharyya (2011) 参照) の三つの規準などにより測定されうる。このほか、Abdallah et al. (2016) が指摘する不正会計検知の研究上の四つの課題、すなわち (a) コンセプトドリフト、(b) 分布の歪み、(c) 巨大なデータ量 (の縮減)、および (iv) リアルタイム検知のサポート、さらには本稿第2節・第3節の各所にまとめられた指摘についても、現在に至るまでの研究の到達点や将来展開の可能性を知るうえで有用なものであると考える。

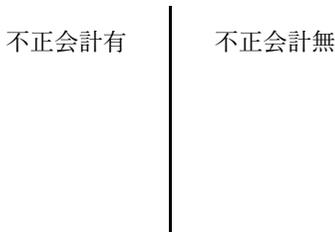


図4. 不正会計の有無の検知 (2区分)

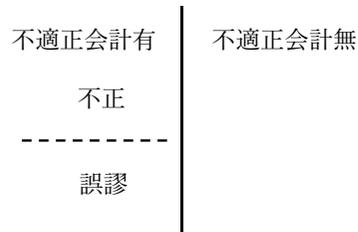


図5. 不適正会計の検知 (3区分)

本(第4)節ではさらに、本稿のレビュー中になした考察から導かれた、関連分野の研究に関して気になった事項について書き留めておきたい。第一に、記述や予測、処方の対象となる現世の事象に関する定義付けや理解、経済的影響の把握、および理論・モデル化が必要であることはいうまでもない。本稿でみた不正会計検知でいえば、経済的影響の甚大さは社会的に有用な研究の要件の一つと考えられ、たとえば Kirkos et al. (2007, 995) や Wang and Yang (2009, 4) といった初期段階から、こうした研究を企図する意義として根底にあったといえよう。また、Humpherys et al. (2011) や Glancy and Yadav (2011) などにおける経営者行動の理論やいわゆる不正のトライアングル²⁵ (Cressay 1953) の研究、さらには会計不正の個別事例の検討 (たとえば、吉見編著 (2018) 参照) についても、「新たな」分析手法の進展とともに却ってその重要性が増しているように思われる。すなわち、ブラックボックス的なモデルを導く技法が一般的となってきたとされる (Amani and Fadlalla 2017, 48) 現状において、分析のインプットとして適切なデータの特性を示し、分析の結果すなわち現世の事象が機能する過程の説明を供するのに、これら研究は資するかもしれない。たとえ

ば、矢口 (2019) は、AI (機械学習) を用いた不正会計予測モデルの活用に関し、「現場では、計算結果の解釈や説明可能性が課題となっている」(122 頁) とし、監査実務に普及するための要件として「監査基準における監査証拠と統計的手法を用いた分析結果が示す証拠の関係整理が不可欠である」(123 頁) と述べる。さらに、「もしかすると」と断ったうえで指摘するが、**図 4** のような現状の不正会計の有無の二分から、**図 5** のように行為者の悪意の有無による不適正会計の不正・誤謬への二分を含む 3 区分 (Amani and Fadlalla 2017, 52)、あるいは不正会計の程度やコストの大小によるさらなる区分 (たとえば、Gepp et al. (2018) と上枝 (2018, 32 頁) 参照) へと向かう研究の展開のためには、統計的・計量的手法の技法の進化と歩調を合わせて協働する別のアプローチが要求されるかもしれない²⁶。第二に、実務家と研究者の関心の差異であり、それは矢口 (2019) を先に引用したように、研究者が分析手法のパフォーマンスの向上に拘泥する傾向があるとしても、実務家にとっては、それに加えて分析結果の解釈や説明可能性、さらには分析ツールの実務への適用可能性が重要となろう。ここでも、先にみた理論的・事例的な研究がその役割を果たすほか、データのビジュアル化 (Dilla and Raschke 2015) も効果的であると考えられる。第三に、適用する手法の特性による分析結果のブラックボックス化の功罪が議論されるべきかもしれない。すなわち、ブラックボックス化は、一方で解釈や説明可能性を低めることは疑いない²⁷ものの、他方では不正会計の隠蔽のための不正者による意図的な偽装工作を妨げるためには役立つかもしれない。第四かつ最後に、意図的に一般に認められた会計基準 (GAAP) を侵犯したり、開示が強制される事項を会計報告書に記載しなかったりすることが不正会計に該当するとするならば、開示基準や会計基準の変化とともに不正会計の定義も変わり、本稿でみた Abdallah et al. (2016) のコンセプトドリフトと同様の問題が生じうると考えられる。当該問題の深刻さについても俄かにはこれを論じえないものの、会計関連の情報を通じたコミュニケーションを考えるわれわれにとっては、念頭に置いておきたい論点といえる。すなわち、何を伝えるようとし、あるいは何を隠蔽しようとし、また何が伝わるのかという視点²⁸を忘れないように心がけつつ、今般レビューした「新たな」分析手法を捉えることも重要である。

最後に、今般のレビューの途上、わが国企業の会計テキストを対象としたさまざまなアプローチの研究が実施されてきていることを了解した (たとえば、伊藤 (2016)、近藤・石光 (2016)、澤登 (2010, 2011, 2013)、廣川 (2013)、および廣瀬 (2017a/b) など)。本稿第 1 節の**図 1** でみたように、企業の実態や将来の帰結が会計処理の原理・原則ないし基準、さらには書き手の意図や選好を通じて会計関連の報告書の上に写像・マッピングされるならば、逆に、会計報告書の適切な読解・分析によると企業の実態や将来の帰結が顕れるとしても、それは決して空想とまではいえない。不正会計に関しては、会計数値という量的な情報が対象となるため、会計報告書のテキスト部分のような質的な情報が不正の存否を顕示する可能性や過程の検討は、個人的には、特に興味の惹かれる研究論点の一つである。何れにしても、本稿執筆中に触れた文献のうち、「新たな」分析手法の将来可能性を否定的にみるものは、筆者の知る限りにおいて存在しなかった。先行する研究者らに伍せるよう、われわれなりの研究を継続していきたい。

【謝辞】 本稿は、「青山学院大学経営学会研究助成金」制度による2018年度の研究助成を受けた成果の一部である。ここに記し、大いに感謝するものである。

注

- 1 彼女たち (Purda and Skillicorn 2015, 1197) によれば、正確には、「財務上の虚偽表示 (financial misrepresentation)」と表記される。なお、財務諸表不正 (financial statement fraud) という表現もみられ、たとえば Tangod and Kulkarni (2015) によれば、「1) 財務報告の操作、2) 財務諸表の作成のもとになる事象、勘定ないしその他重大情報の意図的な省略、あるいは、3) ビジネス取引を測定、認識、報告かつ開示するために用いられる会計原則、政策、および手続の誤った適用」(Ibid., 549) を伴うことがある。なお、本稿において、不正会計および会計不正は、相互に交換可能な用語として区別することなく用いている。
- 2 企業の実態に関しては、不正会計ないし財務上の不正 (financial fraud) の存否以外にも、たとえば、財務的な困窮 (financial distress) の度合いや将来的な倒産の発生、さらには現在の株価の妥当性の判定や株価の推移の予測 (stock market prediction) などが挙げられる (Gepp et al. 2018, 102)。
- 3 ビッグデータは、三つの V、すなわち巨大なデータ容量 (high volume)、高いデータ処理速度 (high velocity) および高いデータの多様性 (high variety)、あるいはこれに加えて高い正確性 (high veracity) と高い価値 (high value) という五つの V という特性を有する (Gartner 2014; Merritt-Holmes 2016)。Gepp et al. (2018, 103) は、これらのうち四つを挙げ、高いデータ容量 (volume) は「あまりに巨大なため伝統的なツールでは不十分なデータ集合」、高いデータの多様性 (variety) は「イメージ、ビデオやその他のフォーマットはもちろん、質的な、テキスト・ベースと混合形式のような、さまざまなデータ・フォーマット」、高いデータ処理速度 (velocity) は「しばしば非常に速いレートで増大している、新規のデータが利用可能になる頻度」、さらに高い正確性 (veracity) は「期を経て劇的に変化しうるデータの質とレlevance」であるとそれぞれ定義する。なお、従来のデータ分析とビッグデータ分析との差異については、以下表注 1 (Davenport 2014, 邦訳 24 頁 = 表 1-1) が参考になる。

概要	ビッグデータ分析	従来のデータ分析
データの種類	非構造化データ	構造化データ
データ量	100 テラバイトから数ペタバイト	数十テラバイト以下
データの流入	常にデータが流入	データは静的な状態
分析手法	機械学習	仮説検証
主な目的	製品やサービスとしての展開	社内での意思決定支援

表注 1. ビッグデータ分析と従来のデータ分析 (出典: Davenport 2014, 邦訳 24 頁 = 表 1-1)

- 4 West and Bhattacharya (2016, 48) によれば、データマイニング (data mining) とは「根源的な意味を導くために大量のデータを処理する任意の方法」をいい、統計的 (statistical) および計算的 (computational) という二つのカテゴリーを有し、このうち前者の統計的なテクニックとは「ロジスティック回帰とベイジアン理論のような、伝統的な数学的手法に基づくもの」であり、後者の計算的な手法とは「ニューラルネットワークとサポートベクトルマシンのような、現代的な人工知能技術を用いるもの」と定義される。両者には多くの類似性もみられるものの、彼らは、統計的な手法がより硬直的なものであるのに対し、計算的な手法は問題領域 (problem domain) から学習をして、それに適応できるという違いがあるとする (Ibid., 48)。
- 5 意思決定有用性という会計 (情報) の変わることのない基本目的に改めて気付かせてくれたのは、Janvrin and Watson (2017, 3) の記述による。
- 6 次 (第 2) 節でみる Amani and Fadlalla (2017, 32) における同論文 (Elliott 1992) の引用は、当該事実、すなわち IT 化の進展はなにも今に始まったことではないことに改めて気付かされる契機となった。
- 7 これらのほか、Richhariya et al. (2012) は、金融不正検知の方法論の包括的かつ簡潔なサーベイを供している。

- 8 FBI (Federal Bureau of Investigation, 連邦捜査局) による定義を用いて彼らは紹介しており、会計不正のほか、企業内部者による自己取引、および会計不正やインサイダー取引の隠蔽を企図した正義 (justice) の妨害のそれぞれが企業不正として調査対象となる活動とされる (Ibid., 562)。
- 9 原著を辿ったところ、Gray and Debreceeny (2011, 379) の参考文献一覧に記載された当該論文タイトルに誤りがあったため、本稿末尾の参考文献表においては正確な表記に改めている。
- 10 彼らの論文の参考文献表 (Abdallah et al. 2016, 112) には、当該箇所の表記に誤りがあるものと思われ、本稿末尾の参考文献表においては修正を施している。
- 11 それぞれ四つの課題のうち、(i) コンセプトドリフトとは、インプットたるデータとターゲットとなる変数の間の関係が期を経て変化する現象をいい (Amani and Fadlalla 2016, 94)、(ii) 分布の歪みは、通常の (不正のない) 事例に比べて不正のある事例が極端に少ないという両クラスに属する標本数に不均衡があることをいい (Ibid., 95)、(iii) 巨大なデータ量 (の縮減) は、不正なデータの次元や特徴・属性・インプット・変数の数が多すぎることにより、データマイニングのプロセスや検知自体が非常に複雑になってしまう状況をいい (Ibid., 96)、さらに (iv) リアルタイム検知のサポートは、システムをオンラインでつなぎ不正の発生直後の検知を可能とするアプローチ (Ibid., 97) のことをいう。
- 12 上枝 (2018) でみた Cecchini et al. (2010)、さらに本稿執筆に際してみた Ravisankar et al. (2011) や Glancy and Yadav (2011) が関連文献とされる。
- 13 宋 (2018, 11-14 頁) は、West and Bhattacharya (2016) の優れた紹介であり、特にその表 2 (同 11-13 頁) は著者 (宋先生) の理解も交えた不正会計検知 (峻別) 手法の特徴、長所および短所の一覧表となっている。
- 14 Bhattacharyya (2011, 608) では、これら規準ないし指標が一般的に用いられていること、すなわち裏を返せば決して新奇のものではないことが明記されている。
- 15 ナラティブ開示の分析に関する彼女たちの簡潔なレビューは、財務会計におけるデータマイニングを議論する Amani and Fadlalla (2017, Sec. 5.3.2., 41) 参照。
- 16 CRISP-DM (Cross Industry Standard Protocol — Data Mining) については、熊野 (2001) 参照。なお、Amani and Fadlalla (2017) の結論部においても、「実務上の推奨事項を明示した、著名な CRISP-DM のフレームワークのレンズを通した会計学のデータマイニングの調査は、適切かもしれない」(Ibid., 52-53) とされ、適用に際して考慮すべき七つの項目が指摘されている。
- 17 彼女たちは、データマイニングが適用される財務会計のサブトピックとして、会計上の概念 (適用数 1)、会計データの質 (適用数 1)、会計基準 (適用数 2)、財務分析 (適用数 6) および財務パフォーマンス (適用数 10) を挙げている。
- 18 監査におけるビッグデータの過少利用ないし利用の進展の遅れの理由のありうべき説明として、彼らは、「監査人は、自身のクライアントが適用するものよりも遥かに先行するような技法やテクノロジーを利用することに気に進まない」(Gepp et al. 2018, 102) ことを挙げる。そして、「ランダムサンプリングという監査技法の初期の利用は、クライアント企業の実務にずっと先行したことに鑑みれば」(Ibid., 102)、当該状況は驚くべきことであるという興味深い指摘をしている。
- 19 前拙稿 (上枝 2018, 32 頁) において指摘したところであるが、金融不正には程度の問題がつきまとい、重大な不正ほど一般的には関係者に及ぼす負の影響、すなわちここでいうところの「金融不正のコスト」が大きいと考えられる。
- 20 たとえば、ある倒産予測モデルにしたがうと将来ある企業が破綻する可能性が合理的に高い水準であることを理由として、継続企業の前提に関する注記が必要になると考え、会社にその旨の開示を要求できるだろうか。会社側が自ら主張する合理的な根拠を盾に開示を拒んだ場合、監査人がとるべき対応はどのようなものだろうか。
- 21 したがって、監査人が自らの職務を奪われる可能性を忌避してビッグデータ利用に二の足を踏んでいるのならば、それは事実ではない、すなわち誤解が生じているかもしれない (Ibid., 108)。
- 22 ちなみに、次数中心性 (degree centrality) と媒介中心性 (betweenness centrality) の二つの基準によって測定されたもっとも影響力のある論文には、上枝 (2018, 18-19 頁) でもみた Loughran and McDonald (2011) が含まれる (Ibid., Figure 4, 7 of 27, and 7-8 of 27)。なお、次数中心性は引用のネットワーク内のノードが有する直接的な結び付きの数で測定され、媒介中心性は他のノードとの間の最短

の経路上にあるノードが存在する程度を意味する (Ibid., 7 of 27)。

- 23 原著では、「Content-Based Criteria Analysis (CBCA)」と表記されていた (たとえば, Humpherys et al. 2011, 586) ものの、誤植であろうと推測し、筆者 (上枝) の理解により修正を施した。
- 24 上枝 (2018, 21 頁) もまた参照。
- 25 今般は深くは考察しないものの、不正のトライアングルの理論に対して批判的な検討をしている Lokanan (2015) は参考になる議論の一つである。同論文の表 1.2 (Ibid., Table 1.2, 219) は、不正のトライアングル (fraud triangle) の欠点をまとめている。
- 26 大橋 (2018) もまた参照。
- 27 もっとも、理論や根拠はなくとも、「鼠をたくさん捕る猫が良い猫である」と考えるむきもあろう。
- 28 たとえば、コミュニケーションだけを考えた場合、確実度の高い不正会計の兆候は、逆に、利害関係者にとって企業の真の姿 (実態) を意図せずとしてではあるが表出しているともいえる。

【参考文献】

- Abdallah, A., M. A. Maarof, and A. Zainal. 2016. Fraud Detection System: A Survey. *Journal of Network and Computer Applications* 68, 90-113.
- Amani, F. A., and A. M. Fadlalla. 2017. Data Mining Applications in Accounting: A Review of the Literature and Organizing Framework. *International Journal of Accounting Information Systems* 24, 32-58.
- Bach, M. P., Z. Krstić, S. Seljan, and L. Turulja. 2019. Text Mining for Big Data Analysis in Financial Sector: A Literature Review. *Sustainability* 11 (5), 1277.
- Bhattacharyya, J., S. Jha, K. Tharakunnel, and J. G. Westland. 2011. Data Mining for Credit Card Fraud: A Comparative Study. *Decision Support Systems* 50, 602-613.
- Bond, C. F., and B. M. DePaulo. 2006. Accuracy of Deception Judgments. *Personality and Social Psychology Review* 10 (3), 214-234.
- Cecchini, M., H. Aytug, G. J. Koehler, and P. Pathalk. 2010. Making Words Work: Using Financial Text as a Predictor of Financial Events. *Decision Support Systems* 50 (1), 164-175.
- Chen, S. 2016. Detection of Fraudulent Financial Statements Using the Hybrid Data Mining Approach. *SpringerPlus* 5: 89, 1-16 of 16.
- Chen, Yuh-Jen., C. Wu, Yuh-Min Chen, H. Li, and H. Chen. 2017. Enhancement of Fraud Detection for Narratives in Annual Reports. *International Journal of Accounting Information Systems* 26, 32-45.
- Cressey, D. R. 1953. *Other People's Money: A Study in the Social Psychology of Embezzlement*. Glencoe, Illinois, The Free Press.
- Davenport, T. H. 2014. *Big Data at Work*. (有限責任監査法人トーマツデロイトアナリティクス監修・小林啓倫訳 (2014), 『データ・アナリティクス 3.0 —ビッグデータ超先進企業の挑戦—』, 日経 BP 社。)
- Dilla, W. N., and R. L. Raschke. 2015. Data Visualization for Fraud Detection: Practice Implications and a Call for Future Research. *International Journal of Accounting Information Systems* 16, 1-22.
- Dong, W., S. Liao, B. Fang, X. Cheng, C. Zhu, and W. Fan. 2014. The Detection of Fraudulent Financial Statements: An Integrated Language Model Approach. *PACIS 2014 Proceedings* 383.
- Elliott, R. K. 1992. Commentary on *The Third Wave Breaks on the Shores of Accounting*. *Accounting Horizons* 6 (2), 61-85.
- Fanning, K. M., and K. O. Cogger. 1998. Neural Network Detection of Management Fraud Using Published Financial Data. *International Journal of Intelligent Systems in Accounting, Finance & Management* 7 (1), 21-41.
- Gao, L., R. P. Srivastiva. 2011. The Anatomy of Management Fraud Schemes: Analyses and Implications. *Indian Accounting Review* 15 (1), 1-23.
- Gartner. 2016. IT Glossary. Available on-line at: <<https://www.gartner.com/it-glossary/big-data/>>. (2019年3月13日アクセス)
- Gepp, A., M. K. Linnenluecke, T. J. O'Neill, and T. Smith. 2018. Big Data Techniques in Auditing Research and Practice: Current Trends and Future Opportunities. *Journal of Accounting Literature* 40, 102-115.

- Glancy, F. H., and S. B. Yadav. 2011. A Computational Model for Financial Fraud Detection. *Decision Support Systems* 50, 595-601.
- Gray, G. L., and R. S. Debreceeny. 2014. A Taxonomy to Guide Research on the Application of Data Mining to Fraud Detection in Financial Statement Audits. *International Journal of Accounting Information Systems* 15, 357-380.
- Hajek, P., and R. Henriques. 2017. Mining Corporate Annual Reports for Intelligent Detection of Financial Statement Fraud — A Comparative Study of Machine Learning Methods. *Knowledge-Based Systems* 128, 139-152.
- Huang, S. Y. 2013. Fraud Detection Model by Using Support Vector Machine Techniques. *JDCTA: International Journal of Digital Content Technology and Its Applications* 7 (2), 32-42.
- Humpherys, S. K., K. C. Moffitt, M. B. Burns, J. K. Burgoon, and W. F. Felix. 2011. Identification of Fraudulent Financial Statements Using Linguistic Credibility Analysis. *Decision Support Systems* 50 (3), 585-594.
- Janvrin, D. J., and M. W. Watson. 2017. “Big Data”: A New Twist to Accounting. *Journal of Accounting Education* 38, 3-8.
- Kennedy, K. A. 2012. An Analysis of Fraud: Causes, Prevention, and Notable Cases. Honors Theses and Capstones. 100. (<https://scholars.unh.edu/cgi/viewcontent.cgi?article=1099&context=honors>) (2019年3月4日アクセス)
- Kirkos, E., C. Spathis, and Y. Manolopoulos. 2007. Data Mining Techniques for the Detection of Fraudulent Financial Statements. *Expert Systems with Applications* 32, 995-1003.
- Li, F. 2010a. The Information Content of Forward-Looking Statements in Corporate Filings — A Naïve Bayesian Machine Learning Approach. *Journal of Accounting Research* 45 (5), 1049-1102.
- Li, F. 2010b. Textual Analysis of Corporate Disclosures: A Survey of Literature. *Journal of Accounting Literature* 29 (1), 143-165.
- Li, J., K. Huang, J. Jin, and J. Shi. 2008. A Survey on Statistical Methods for Health Care Fraud Detection. *Health Care Management Science* 11 (3), 275-287.
- Lokanan, M. E. 2015. Challenges to the Fraud Triangle: Questions on Its Usefulness. *Accounting Forum* 39 (3), 201-224.
- Loughran, T., and B. McDonald. 2011. When is a Liability not a Liability? Text Analysis, Dictionaries, and 10-Ks. *Journal of Finance* 66 (1), 35-65.
- Merritt-Holmes, M. 2016. Big Data & Analytics: The DNA to a Successful Implementation in 2016. Available on-line at: (<https://dataanalytics.report/view-resource.aspx?id=2892>). (2019年3月13日アクセス)
- Ngai, E. W. T., Y. Hu, Y. H. Wong, Y. Chen, and X. Sun. 2011. The Application of Data Mining Techniques in Financial Fraud Detection: A Classification Framework and an Academic Review of Literature. *Decision Support Systems* 50 (3), 559-569.
- Purda, L., and D. Skillicorn. 2015. Accounting Variables, Deception, and a Bag of Words: Assessing the Tools of Fraud Detection. *Contemporary Accounting Research* 3 (3), 1193-1223.
- Ravisankar, P., V. Ravi, G. R. Gao, and I. Bose. 2011. Detection of Financial Statement Fraud and Feature Selection Using Data Mining Techniques. *Decision Support Systems* 50, 491-500.
- Richhariya, P., P. K. Singh, and E. Duneja. 2012. A Survey on Financial Fraud Detection Methodologies. *IRACST — International Journal of Commerce, Business and Management (IJCBM)* 1 (1), 14-24.
- Segal, S. Y. 2016. Accounting Frauds — Review of Advanced Technologies to Detect and Prevent Frauds. *Economics and Business Review* 2 (4), 45-64.
- Sharma, A., and P. K. Panigrahi. 2012. A Review of Financial Accounting Fraud Detection Based on Data Mining Techniques. *International Journal of Computer Applications* 39 (1), 37-47.
- Tangod, K. K., and G. H. Kulkarni. 2015. Detection of Financial Statement Fraud Using Data Mining Technique and Performance Analysis. *International Journal of Advanced Research in Computer and Communication Engineering* 4 (7), 549-555.

- Van den Bogaerd, M., and W. Aerts. 2011. Applying Machine Learning in Accounting Research. *Expert Systems with Applications* 38 (10), 13414-13424.
- Wang, J., and J. G. S. Yang. 2009. Data Mining Techniques for Auditing Attest Function and Fraud Detection. *Journal of Forensic & Investigative Accounting* 1 (1), 1-24.
- West, J., and M. Bhattacharya. 2016. Intelligent Financial Fraud Detection: A Comprehensive Review. *Computers & Security* 57, 47-66.
- Zaki, M., and B. Theodoulidis. 2013. Analyzing Financial Fraud Cases Using a Linguistics-Based Text Mining Approach. Available at SSRN 2353834.
- Zhou, W., and G. Kapoor. 2011. Detecting Evolutionary Financial Statement Fraud 50, 570-575.
- 伊藤健頭 (2016), 「SEC 基準採用企業の MD&A 情報と株式市場の反応」, 『年報経営ディスクロージャー研究』, 第 14 号 (2016 年 3 月), 53-62 頁。
- 上枝正幸 (2014a), 「会計学のコミュニケーション研究に向けて—コミュニケーション学/論の援用可能性の探索—」, 『青山経営論集』, 第 49 巻第 2 号, 17-37 頁。
- 上枝正幸 (2014b), 「(解題深書) 企業の情報開示 (ディスクロージャー) と会計コミュニケーション研究」, 『企業会計』, 第 66 巻第 12 号, 89-93 頁。
- 上枝正幸 (2014c), 「わが国の会計コミュニケーション研究の回顧と展望」, 『青山経営論集』, 第 49 巻第 3 号, 94-111 頁。
- 上枝正幸 (2016a), 「会計コミュニケーションを研究する—過去・現在・未来—」, 『青山経営論集』, 第 50 巻第 4 号, 23-40 頁。
- 上枝正幸 (2016b), 「裁量的会計ナラティブ—経営者の言説の研究—」, 『青山経営論集』, 第 51 巻第 3 号, 169-192 頁。
- 上枝正幸 (2018), 「企業の実態はいかにして表出するのか—会計ディスクロージャーの統計的言語分析の可能性—」, 『青山経営論集』, 第 53 巻第 2 号, 15-37 頁。
- 大城直人 (2014), 「不正会計の早期発見に関する海外調査・研究報告書」, 金融庁金融研究センター, DP2014-6 (2014 年 8 月)。
- 大橋和彦 (2019), 「機械学習の経済分析への利用—近年の発展」, 『証券アナリストジャーナル』, 第 57 巻第 1 号 (2019 年 1 月号), 61-65 頁。
- 熊野匡 (2001), 「データマイニング・プロジェクト・ライフサイクル—データマイニングと CRISP-DM 方法論の紹介—」, 『SOFTECHS』, 第 24 巻第 1 号 (2001 年 6 月発行), 76-88 頁。
- 近藤隆史・石光裕 (2016), 「経営者のマネジメント・コントロールの意識と企業業績—有価証券報告書のテキスト分析を通して—」, Discussion Paper Series, No. 2016-01, 1-21 頁。
- 澤登千恵 (2010), 「テキストマイニングを活用した財務報告実務の展開の考察」, 『研究紀要 (高松大学・高松短期大学)』, 第 52・53 合併号, 39-70 頁。
- 澤登千恵 (2011), 「19 世紀イギリス鉄道会社におけるコストマネジメントの背景—テキストマイニングによる分析結果の会社間比較を中心に—」, 『研究紀要 (高松大学・高松短期大学)』, 第 54・55 合併号, 179-207 頁。
- 澤登千恵 (2013), 「テキストマイニングを活用した Great Western 鉄道会社の会計変化の検討」, 『大阪産業大学経営論集』, 第 15 巻第 1 号, 27-42 頁。
- 首藤昭信 (2019), 「AI が会計学研究に与える影響」, 『會計』, 第 195 巻第 2 号 (2 月号), 127-141 頁。
- 宋明子 (2018), 「昨今の経済環境等の変化に対応した不正会計の早期発見に関する調査・研究」, 金融庁金融研究センター, DP2017-6 (2018 年 2 月)。
- 瀧博・井上直樹・小俣光文・笠井直樹・小松義明・坂上学・高原利栄子・田口聡志・中村映美・藤岡英治・堀古秀徳 (2018), 『テクノロジーの進化と監査』, 日本監査研究学会・課題別研究部会, 平成 30 年度中間報告 (2018 年 8 月, 日本監査研究学会第 41 回全国大会 (於: 近畿大学))。
- 日本公認会計士協会 (2016), 「IT を利用した監査の展望~未来の監査へのアプローチ~」, IT 委員会研究報告第 48 号。
- 日本公認会計士協会 (2019), 「次世代の監査への展望と課題」, IT 委員会研究報告第 52 号。
- 廣川佐千男 (2013), 「文単位の有価証券報告書分析による利益伸び率の予測」, 『信学技法 (電子情報通信学

会)』, Vol. 113, No. 213, 77-82 頁。

廣瀬喜貴 (2017a), 「公会計情報の言語的特徴: 東京、ニューヨーク、ロンドンを対象とした分析」, 『高崎経済大学論集』, 第 59 卷第 2・3・4 合併号, 61-73 頁。

廣瀬喜貴 (2017b), 「可読可能性研究の課題と展望」, 『高崎経済大学論集』, 第 60 卷第 2・3 号, 75-87 頁。

矢口龍一 (2019), 「会計における AI (IT) の現状」, 『會計』, 第 195 卷第 2 号 (2 月号), 113-126 頁。

吉見宏編著 (2018), 『会計不正事例と監査 (日本監査研究学会リサーチ・シリーズ XVI)』, 同文館出版。

『日本経済新聞』, 2018 年 10 月 8 日 (月), 朝刊, 第 12 版, 11 面, 「監査 AI 生かす」。