



RIETI Discussion Paper Series 19-J-039

機械学習手法を用いた不正会計の検知と予測

宇宿 哲平
あずさ監査法人

近藤 聡
あずさ監査法人

白木 研吾
あずさ監査法人

菅 美希
KPMG LLP

宮川 大介
一橋大学



Research Institute of Economy, Trade & Industry, IAA

独立行政法人経済産業研究所
<https://www.rieti.go.jp/jp/>

機械学習手法を用いた不正会計の検知と予測*

宇宿哲平（あずさ監査法人）、近藤聡（あずさ監査法人）、白木研吾（あずさ監査法人）、
菅美希（KPMG LLP）、宮川大介（一橋大学）**

要 旨

本研究は、不正会計の検知と予測における機械学習手法の有用性を実証的に検討したものである。第一に、既存研究における限定的な変数に基づくパラメトリックモデルに代えて機械学習手法を用いることで、高次元の変数に基づく不正会計の「検知」を行い、その精度の改善を検証する。第二に、既存研究において十分に検討されていない不正会計の「予測」を同様のフレームワークで行う。第三に、構築された予測モデルの解釈を行う目的から、不正会計の予測結果（スコア）が、どのような情報に反応しているかを検討する。本邦上場企業の不正会計イベントを対象とした実証分析から、機械学習手法の利用が検知精度の改善に一部貢献するほか、当該手法の利用によって参照可能となる高次元の変数利用が検知精度の大幅な改善に貢献していることを確認した。また、機械学習ベースのモデルが実務的観点から十分な水準の予測精度を実現できることも確認した。本研究では、既存研究が理論的な想定に基づいて参照してきた変数以外の変数（例：ガバナンス関連変数）について、不正会計の検知・予測にあたって一定程度有用な情報が含まれていることも確認しており、不正会計の発生メカニズムに関する今後の理論的検討において参照すべき実証事実を示すものと考えられる。

キーワード：不正会計、機械学習、不正検知、不正予測、解釈可能性

JEL classification: M42, C53, C14

RIETI ディスカッション・ペーパーは、専門論文の形式でまとめられた研究成果を公開し、活発な議論を喚起することを目的としています。論文に述べられている見解は執筆者個人の責任で発表するものであり、所属する組織及び（独）経済産業研究所としての見解を示すものではありません。

*本稿は、国立大学法人一橋大学と有限責任 あずさ監査法人との共同研究および独立行政法人経済産業研究所におけるプロジェクト「企業金融・企業行動ダイナミクス研究会」の成果である。本稿の原案に対して、大橋和彦（一橋大学）、本多俊毅（一橋大学）、中川秀敏（一橋大学）、鈴木健嗣（一橋大学）、植杉威一郎（一橋大学）、小野有人（中央大学）、細野薫（学習院大学）、小川勤（あずさ監査法人）、清水多賀雄（あずさ監査法人）、深見英二（あずさ監査法人）、宮下淳（あずさ監査法人）、中島厚志（経済産業研究所）、矢野誠（経済産業研究所）、森川正之（経済産業研究所）、ならびに経済産業研究所ディスカッション・ペーパー検討会の方々から多くの有益なコメントを頂いた。ここに記して、感謝の意を表したい。

** 宇宿、近藤、白木：有限責任 あずさ監査法人 Digital Innovation 部 〒100-0004 東京都千代田区大手町 1-6-1 大手町ビルディング7階、菅：KPMG LLP Financial Due Diligence 1350 Avenue of the Americas, New York, NY, UNITED STATES, 10019、宮川：Corresponding author、一橋大学大学院経営管理研究科 准教授 〒101-8439 東京都千代田区一ツ橋 2-1-2 E-mail: dmiyakawa@hub.hit-u.ac.jp.

1. はじめに

国内外における会計不祥事の実例を参照するまでもなく、企業の財務情報に関するミスレポーターティング（いわゆる「財務諸表における虚偽表示」）は実務的観点から回避すべき重要な経済インシデントといえる。特に、上場企業における財務情報のミスレポーターティングは、株式市場、債券市場、銀行取引などの金融取引に関与する様々な経済主体の意思決定を歪め、非効率な資源配分をもたらす。また、こうしたミスレポーターティングは、実物取引関係にある販売先や仕入先にとっても認識しない過大なリスクを引き受けさせかねない。結果として、当該リスクが顕在化した際に、意図しない経済活動の停滞が生じる可能性がある。より深刻な事態として、こうしたミスレポーターティングが報告者の意図としてなされる場合（いわゆる「不正会計」）が頻繁に生じる場合において、金融・実物取引市場が正常に機能しない可能性も想像されるだろう。

こうした財務情報のミスレポーターティング、特に不正会計に起因する問題は事業活動に限られたものではなく、政策運営にも及ぶ。例えば、企業の財務情報は諸々の政策介入の際に参照されているため、外部から観察した企業の属性が実態と乖離している場合、意図した政策が正しく実施されない可能性がある。また、政策効果を事後的に評価する場合にも、こうした不正会計の存在は計測結果に重大な誤りをもたらさう。

こうした問題意識を基に、主として会計学の領域において、不正会計の発生メカニズムに関する理論的研究が進められてきた（例：Dechow et al. 1996）。こうした理論的研究の第一の類型は、企業が不正会計を行う「理由」に関するものである。業績低迷に直面した企業が、資金調達に際して財務情報を意図的に改ざんすることで、より良い調達条件を実現しようとするといったメカニズムは、こうした理論的検討の一例である。これらの理論的検討は、何らかの目的関数と制約条件を所与とした場合における企業の最適化行動の有様を記述することで、不正会計に繋がる決定要因を特定したものと整理できる。

理論的研究の第二の類型は、会計分野における実務的な知見を参照する形で、不正会計の発生と「相関」を有する財務情報を特定しようとするものである。例えば、会計上の利益から実際のキャッシュフローを取り除いた「会計発生高」のうち、一定の仮定の下で正常に発生すると考えられる額（非裁量的会計発生高）を控除した結果として計測される額を「裁量的会計発生高」と呼ぶ。理論的研究の第二の類型では、経営者の利益調整行動に連動していると考えられるこの裁量的会計発生高が観察される場合において、何らかの不正会計が発生している可能性が高いという推論を行っている。

これらの理論的な検討を踏まえて、近年では、不正会計を行っている可能性が高い企業の判別（各期の財務諸表に関する不正会計の検知）を目的とする実証的取り組みが進められている。具体的には、不正会計の発生を検知するための統計モデルの構築が進められており、代表的な先行研究として、Dechow et al. (2011)や Song et al. (2016)が存在している。これらの実証研究では、上記の二つの方向性における理論的な議論を踏まえて、検知に当たって

参照する変数の選択をマニュアルで行った上で、パラメトリックなモデルを用いて、現時点における不正会計の発生がそれらの変数とどの様に関係しているかを推定した上で、不正会計の検知を試みている。これらの既存研究は、理論的な検討と整合的な形で、検知に際して参照した変数が機能していることを確認すると共に、主としてパラメータ推定に用いたデータ内での良好な当てはまりを報告している。

このように、理論的検討と整合的な結果を報告する実証研究が蓄積されている一方で、実証分析の観点からは少なくとも二つの課題を指摘することができる。まず、第一の課題として、現時点における不正会計の検知を目的とした取り組みにおいて、検知に際して参照する変数として活用されていない膨大な変数が存在する。例えば、Song et al. (2016)では、理論的な想定に基づいて合計6種類の変数をモデルに投入しているが、分析対象企業を取り巻く膨大な情報の中には、これらの限定的な変数を検知の精度改善という意味で補完するものが含まれている可能性がある。こうした制約的な変数選択は、理論的想定をテストする目的には適っているものの、検知を目的とした取り組みにおいては改善の余地を残している。勿論、高次元の変数を含めたモデル構築に関して、既存研究が用いてきたパラメトリックモデルでは参照可能な変数の数に限界が存在するため、限定的な種類の変数のみを用いたモデル構築に留まっていることには致し方ない面もあるが、手法面の工夫によってこの点を克服し、より高い検知精度を実現することが期待されている。

本研究の第一の目的は、以上の議論を踏まえて、膨大な説明変数を参照したモデル構築が可能な機械学習手法を用いた不正会計の検知モデルを構築することにある。なお、モデルの構築に当たっては、機械学習手法を用いた不正会計の検知モデル構築を行っている近年の代表的研究である Perols (2011)および Perols et al. (2017)と同様の問題意識から、不正会計イベントの観測数が少ないこと（いわゆる「imbalance problem」）に対応した処理も行う。また、既存研究が参照してきた理論的検討の成果を踏まえつつ、可能な限り幅広い企業情報を説明変数として用いる。こうした追加の変数には、企業の財務情報から構築される多様な財務指標のほか、株主、役員、従業員情報を中心として構築したガバナンス関連変数、取引銀行情報を用いて構築した銀行取引変数が含まれる。

次に、第二の課題として、ほぼ全ての既存研究が、現時点での不正会計の「検知」に分析の主眼を置いており、将来における不正会計の発生「予測」を明示的に取り扱っていないという点が挙げられる。予測分析の典型的な対象である倒産とは異なり、不正会計については現時点における分析対象イベントの発生を完全に観察できないことが多く、「検知」にも実務的に大きな価値がある。しかし、例えば、監査法人の業務上の観点からは、会計上のミスレポーティングが生じる蓋然性の高い企業を事前にスクリーニングすることにも大きな意味があるだろう。

本研究の第二の目的は、機械学習手法ベースの分析フレームワークを、既存研究が十分に取り扱ってこなかった不正会計の予測モデル構築（将来時点における不正会計イベントの

企業レベル予測) に応用した上で、その予測精度を検証することにある¹。

本邦上場企業のデータを用いた分析から得られた不正会計の「検知」精度に関する検証結果は、以下の通り要約される。検知精度検証に当たっては、精度の改善度合いを確認する目的から、既存研究の結果に倣った限定的な変数のみに依拠したパラメトリックモデル(モデル1)、同様に限定的な変数のみに依拠しながらも機械学習手法を用いたノンパラメトリックモデル(モデル2)、そして、変数を拡充した上で機械学習手法を用いて構築したノンパラメトリックモデル(モデル3~16、特にモデル12を全ての変数を投入したフルモデルとして参照する)の検知精度を比較した。

第一に、機械学習手法の利用(モデル1と2の比較)と当該手法の利用によって参照可能となる高次元の変数利用(モデル2と12の比較)の両方が検知精度の改善に貢献していることを確認した。検知精度の評価において一般的に用いられている **hold-out data** (いわゆる「**out-of-sample data**」)における AUC に注目した場合、不正会計の検知に関するモデル1の AUC が 0.71 程度であるのに対して、モデル2の AUC は 0.79 程度であり、モデル12では更に 0.87 程度まで改善する。モデル12における変数の拡充は機械学習手法の利用があって初めて可能となるものであるため、「機械学習手法の活用」と「変数の拡充」が各々不正会計の検知精度向上に対してどの程度の貢献を有しているかを独立に計測することは容易ではないが、上記の結果は、既存モデルと同様の変数群を利用した場合でも、検知モデルの構築手法自体に工夫の余地があることを示唆している。

第二に、予測に当たって参照する情報が観察された時点から向こう一年間の不正会計の発生を対象として、機械学習ベースの不正会計「予測」モデルを構築した上で、その **hold-out data** での予測精度を計測したところ、実務的観点から十分な水準の精度が実現できることを確認した。具体的には、前述の AUC で見た場合、不正会計の予測に関するモデル1の AUC が 0.68 程度に留まった一方で、モデル12では 0.82 程度となっている。この結果は、既存研究がこれまで明示的に取り扱ってこなかった不正会計の将来予測においても、機械学習手法を用いて高次元の変数を用いたモデル構築を行う価値が認められることを意味している。

第三に、これらの検知・予測モデルの構築に当たって、既存研究が理論的な想定に基づいて参照してきた変数以外の変数(例:ガバナンス関連変数)にも一定程度の貢献が認められた。具体的には、従業員の平均勤続年数や役員持ち株比率などが、将来の不正会計の発生予測に対する貢献度 (**importance**) の意味で上位にランクされている。また、これらの変数が平均的な企業属性の近傍で変動した場合に予測スコアがどの様に変動するかを確認したところ、これらの変数の水準が低下するに従って不正会計の発生確率が上昇することが確認された。こうした結果は、既存研究において提示されていた理論的想定が示唆する変数の他にも、検知・予測に対して大きな貢献を有する変数が多く存在する可能性があることを意味

¹ 不正会計を含む金融分野での詐欺(クレジットカード、証券、保険)を対象とした近年の検知・予測モデル開発動向のサーベイとしては、West and Bhattacharya (2016)を参照されたい。

している。この意味で、本稿で得られた結果は、高次の変数を含む検知・予測モデルを実務的に活用する余地が大きいことを示すと共に、不正会計の発生メカニズムに関する今後の理論的検討に当たっても有用な情報を提供しているものと考えられる。

本稿の構成は以下の通りである。第2節では、本研究の前提となる既存研究を概観する。第3節では、本稿での分析の前提となる不正会計の制度的背景を説明する。第4節、第5節では、各々本稿における実証戦略と分析用データについて説明した上で、第6節で実証分析の結果をまとめる。第7節及び第8節では、本稿での分析を踏まえたディスカッションとまとめを示す。

2. 先行研究

会計分野における不正会計の理論的研究には、不正会計の決定要因に関して理論的な検討を行った一連の研究と、不正会計の発生パターンに関して主として会計処理の観点から議論を行ったものがある。前者の一例として、Dechow et al. (1996)では、資金調達に際しての企業の利益計上動機を挙げている。後者の一例としては、利益の質と不正会計の発生との間の相関関係を議論した一連の研究がある（例：Dechow et al. 2010）。

本稿の主たるテーマである、不正会計の検知・予測モデル構築に関する代表的な研究としては、Dechow et al. (2011)が存在する。彼らは、利益の質、パフォーマンス、非金融指標、オフバランス活動、市場関連変数といった、理論的に見て不正会計の発生と関連している可能性の高い変数を選択した上で、標準的なパラメトリックモデルである制限従属変数モデルを推定することで、不正会計の発生と同時点における情報との相関関係を把握し、現時点における不正会計の発生確率に対応するスコアを算出している。同様に、Song et al. (2016)では、Dechow et al. (2011)で参照されている変数に加えて、実物経済活動に関するマニピレーション、会計上の保守主義、日本特有の要因に関連する変数を採用しパラメトリックモデルを推定している。何れの分析においても、理論的な議論から設定された変数が、想定された符号を伴って不正会計の発生スコアと相関していることを確認すると共に、主として推定に用いたデータ内（インサンプル）でのモデルのフィットが高いことを示している。

財務諸表に関するミスレポーターを含む financial fraud 一般について、その検知・予測を対象とした近年の実証研究をサーベイしたものとしては、West and Bhattacharya (2016)が存在する。特に、機械学習手法を用いた不正会計の検知モデル構築を行っている近年の代表的研究である Perols (2011)および Perols et al. (2017)では support vector machine を用いて不正会計の検知を行うと共に、データの imbalance problem への対処についても議論している²。

これらの先行研究の存在を踏まえた本研究の貢献は以下の二点である。第一に、機械学習

² 企業の倒産などを対象として、機械学習ベースのモデルを構築して予測を試みた例は多く存在する（例：Miyakawa 2019a）。

ベースのモデルを用いて、imbalance problem への対応を行いながら財務諸表に関するミスレポーティングの検知と予測を網羅的に行った研究は現状では極めて限られている。本研究の試みは同分野の実証的知見の蓄積に貢献するものと考えられる。第二に、宋(2018)でも議論されている通り、既存研究が参照してきた情報以外にも不正会計の検知・予測に貢献する変数が存在する可能性は高い。本研究では、既存研究が参照してきたもの以外の情報として、会計監査実務で用いられているような指標やガバナンスに関する指標、銀行取引に関する指標を利用することでこの可能性を検証している。

3. 不正会計に関する実務的バックグラウンド

3.1 不正会計の種類

本稿では、不正会計を「財務会計情報について実態と異なる内容を開示する行為」と定義する。過去に金融庁から課徴金の納付命令を受けたケース及び証券取引等監視委員会から有価証券報告書等の虚偽記載に対する告発があったケースを参照すると、こうした定義に該当する不正会計の事例には、幾つかの種類が存在する。

第一に、架空売上の計上などを通じた収益(例:売上高)の過大計上である。損益計算書のトップラインに当たる売上高の過大計上は、費用項目が実態に即して計上されている場合において、当該過大計上額に相当する税引前利益の過大計上に繋がる。過去の事例を見る限り、この類型に該当する不正会計事例が最も多い。第二に、売上原価を中心とする費用の過少計上である。過去の事例では、このほかにも、広告宣伝費等の販売費及び一般管理費に関する過少計上も確認されている。こうした費用の過少計上については、減損損失、貸倒引当金、債務保証損失引当金といった経営者の見積もりを伴う評価性の費用項目に関しても見られる。最後の類型として、第三に、のれんを中心とした資産の過大計上が挙げられる。この類型に該当する資産の過大計上は、たな卸資産のような流動資産、生産設備に代表される有形固定資産、ソフトウェアに代表される無形固定資産が典型的である。資産の過大計上は負債の部を所与とすれば、純資産の部の過大計上に繋がる³。

本研究では、これらの類型全てを一括りとして検知・予測の対象として取り扱うこととし、類型別の検知・予測モデル構築については将来の課題として位置付ける。不正会計の検知及び予測を対象としたモデル構築の文脈では、異なる類型ごとにモデルを構築する必要性も議論されている(例:Perols et al. 2017)。実際に、倒産と休廃業のように、企業の退出という意味では共通のイベントであってもその予測に貢献する変数に大きな差異が生じることや、各イベントの発生に対応するスコアが必ずしも高い相関を有していないことは機械学習ベースのモデルを用いた既存研究からも示唆されている(例:Miyakawa 2019a)。

³ 売上高の過大計上、費用の過少計上、資産の過大計上以外の類型としては、負債の過少計上が挙げられるが、過去の実例で見るとこれはマイナーな類型と言える。

3.2 不正会計に関する制度

本節では、本稿における不正会計イベントの計測に当たって参照する制度を概観する。第一に、本稿では有価証券報告書の虚偽記載等で金融庁による課徴金納付命令を受けた企業を不正会計イベントが生じた企業として認識する。当該制度は、証券市場の公正性・透明性を確保するために、証券市場への信頼を害する違法行為に対する行政上の措置として2005年4月に導入され、違法行為を行った者に対して金銭的負担を課すものである。課徴金納付制度の対象とする違法行為には、インサイダー取引などの不公正取引の他、有価証券報告書等の虚偽記載等（継続開示義務違反等）が含まれ、金融庁長官が課徴金納付制度に基づく課徴金納付命令を行う。第二に、本稿では、証券取引等監視委員会が公開財務諸表の虚偽記載等について、虚偽有価証券報告書提出事件として検察庁に対して告発した企業についても、不正会計イベントが生じた企業として認識する。

上記の虚偽記載等を調査する対象は有価証券報告書等であるが、これらには公衆縦覧期間が存在し有価証券報告書等では5年間である。このため、上記の制度において指摘を受けた虚偽記載に対応する形で訂正を行う場合、その期間は公衆縦覧期間に限られる。後述の通り、本稿では上記の二制度に対応する重大なイベントを検知・予測する実務上の要請を踏まえて不正会計の計測を行うが、代替的な計測方法として、これらの不正会計に対応して生じた付随的な訂正についても不正会計イベントとして計測したケースを分析の対象とする。

なお、本稿が対象とする検知・予測の対象は、あくまで上記の制度によって発覚したイベントのみである。このため、実際には不正会計が存在していたにも関わらず上記制度によっては発覚していないケースについて、検知・予測対象のアウトカムに関する計測上の誤りが存在することになる。ただし、これらのケースの中には、最終的に財務諸表の訂正が行われたケースもあり、この場合は訂正報告書などから捕捉することが可能である。これらの追加的なデータを用いた分析を行うことによって、検知・予測モデルの頑健性を検討することは可能であり、将来の重要な検討課題としたい。

4. 分析手法

本節では、不正会計の検知と予測を目的とする機械学習ベースのモデル構築について、モデルのトレーニングとテストの二つのステップを順に概観する。最も重要な点はモデルの推定（トレーニング）に当たって用いたデータ以外のデータ（hold-out data）を用いて検知・予測の精度評価を行う点にある。

4.1 分析手法：トレーニング

構築すべきモデルの構造は、分析の目的が同時点において発生しているイベントの検知であるか、もしくは将来時点において発生するイベントの予測であるかに依存しない。すなわち、検知・予測時点において入手可能な情報ベクトル \mathbf{X} から、特定のイベントの現時点もし

くは将来時点における発生確率に対応するスコア S を算出することが目的となる。

$$S = F(\mathbf{X})$$

データの観測時点に関して、検知の場合、 X と検知対象イベントが同時点で観測されることを妨げないが、予測の場合、 X は予測対象イベントの観測時点より前に観測されている必要がある。 X がどの程度前もって観測されている必要があるかは、イベントの発生をどの程度前もって予測するかに依存する。例えば、分析時点から向こう一年間のウインドウで不正会計の発生を予測するのであれば、当該一年間のウインドウの始点において分析対象が観察可能な情報が X として用いられる。

会計領域における既存の実証研究の多くは、イベントの観測と非観測を1/0の二値変数として記録し、ロジットやプロビットといった制限従属変数モデルにおいて誤差項 ε における適当な仮定を置いた上で、潜在変数 L^* を構成するパラメータ (α, β) を推定することで、イベントの発生に関する構造をパラメトリックに把握し、その推定結果と個々の企業の情報ベクトル \mathbf{X} を参照することで S を推定するという手順を取っている⁴。こうした手法は、検知・予測に当たって参照すべき情報ベクトル \mathbf{X} が何らかの事情によって事前に明らかになっており、かつ \mathbf{X} の次元がパラメータ (α, β) の推定に際して問題のない次数に留まっている場合において適切な手法となり得る。

$$S = F(\mathbf{X}; \alpha, \beta) = \text{Prob}(L^* > 0) \quad \text{where } L^* = \alpha + \mathbf{X}\beta + \varepsilon$$

こうした手法が検知・予測の文脈で幾つかの問題を有している点に注意が必要である。第一に、検知・予測に当たって参照すべき情報ベクトル \mathbf{X} が、必ずしも明らかになっていない可能性がある。既述の通り、不正会計の検知に係る実証的な取り組みは、会計分野における理論的な議論を下敷きとして参照すべき情報ベクトルの選択を行ってきた。しかし、宋(2018)等において議論されている通り、会計分野における既存の理論研究が十分に注目していないにも関わらず、実務的な観点や経済学的な観点からは情報として参照することが妥当であるものも多く存在する(例:ガバナンス変数)。すなわち、不正会計イベントの検知をより高い精度で行うことを目的とするのであれば、可能な限り高次元の情報をモデルに投入することが期待されるが、ロジットやプロビットといったパラメトリックモデルはこうした目的にはそぐわないという技術的な問題がある。第二に、モデルの構造がプロビットやロジットが想定している非線形性を超えた複雑な様相を示す可能性がある。上記のモデ

⁴ 検知及び予測の実務的な手順としては、(1) こうしたパラメータの推定をあるサンプルを用いて行ったうえで、(2) こうした推定に用いていないサンプル(hold-out data)を用いて検知・予測精度を評価し、十分な精度が確保されていることを確認した上で、(3) 検知予測対象企業の情報 \mathbf{X} と推定されたパラメータを用いて、 S を算出する。

ルが示す通り、大半の制限従属変数モデルは潜在変数に関する線形のモデルと誤差項の分布に関する特定の仮定の下で、 S をアウトカムとする非線形モデルの推定を行う。こうした状況の下で、モデルに複雑な非線形性を導入するためには、 \mathbf{X} の構築に合って複数の情報からなる変数（例：交差項）や特定の情報の非線形項（例：高次項、離散変数に変換）を用いるなどの方策が考えられるが、この結果として、第一の問題として挙げた変数の次元数が再度問題となる可能性が高い。

こうした問題を踏まえて、本研究では、機械学習ベースの検知・予測モデル構築を行う。具体的には、決定木ベースの予測モデルをサブサンプル毎に構築した結果を合算（アンサンブル）する **Random Forest** を用いる。決定木の構築に当たっては、まず、不正会計が観測されたサンプルと観測されていないサンプルが混在している状態から始めて、 \mathbf{X} に含まれる個々の属性に関する閾値を設定する。次に、当該変数・閾値でサンプルを分割した際の **information gain**（観測/非観測サンプルの混在度合いの減少分）を何らかの指標で計測し、この **information gain** が最大となる変数と閾値のペアを特定した上で、分岐のルールを設定する。以下、最終的に観測もしくは非観測の何れかのみが含まれる状態に到達するまでこの手順を繰り返すことで、決定木を構築する。この際、データ全体に対して決定木を構築するのではなく、サンプルの一部や変数の一部を用いて決定木を構築するという手法を **Random Forest** と呼ぶ（Breiman 2001）。分析手法から明らかなように、古典的なパラメトリックモデルにおける情報の次元数に関する問題は、ここでは大きな問題とはならない点に注意されたい。

こうした手法面の工夫の一方で、検知・予測分析の対象である不正会計イベントの特性から、追加的な考慮が必要となる事柄が存在する。具体的には、不正会計の観測が非観測に比して稀であることに配慮する必要がある。こうした問題は、レアイベントの検知・予測が必要となる様々な局面で生じるものであり、例えば、倒産イベントの予測においても同種の考慮が必要となることが一般的である。本研究では、Chen et al. (2004)によって提案された、**Random Forest** の拡張版である **Weighted Random Forest** を用いる。この手法は、個々の決定木構築及びアンサンブルに際してレアイベントへ相対的に大きなウェイトを課するものである。

個々の決定木からなる判別/分類器を以上の手続きで構築するステップを、以下ではトレーニングステップと呼ぶことにする。このステップを完了することで、分析者の手元には \mathbf{X} を投入することで S を返すノンパラメトリックな関数（分類器）が得られる。

4.2 分析手法：テスト

トレーニングステップで構築された分類器を実務で利用する前に、当該分類器がどの程度の検知・予測精度を実現することができるかを、正確に確認する必要がある。こうした目的から、トレーニングステップで利用していないテストデータ（**hold-out data**）に含まれる \mathbf{X} を分類器に投入することで S を算出し、その水準を参照することで不正会計イベントの有無に関する予測を行い、その予測結果をトレーニングデータ内に記録されている実際のイベン

ト有無レコードと比較することで予測の精度を評価する。代表的な方法としては、 S に関する閾値を何らかの基準に従って設定し、accuracy、precision、recall、F-measureなどの精度指標を参照するやり方が用いられる。こうした方法は、精度評価の結果について直観的に理解できるという長所を有している一方で、特定の閾値に依存した評価であるという短所がある⁵。そこで、本稿では、より頑健な精度評価の手法として、ROC 曲線に基づく AUC を参照する。ここで ROC 曲線とは、上記の閾値を変化させたときに (False-positive rate, True-positive rate) からなる座標が描く軌跡であり、この曲線の下面積を参照することで、検知・予測の精度を一つの数字として表現することができる⁶。

4.3 学習済みモデルの“interpretability”

機械学習ベースのノンパラメトリックモデルに対する批判の一つとして、検知・予測の結果 S の変動を、 X 内の特定の情報の変動と関連付けて解釈することが難しい、という指摘がある。確かに、前節でのトレーニングステップに関する説明からも明らかなように、複数の決定木のアンサンブルとして得られる予測結果を特定の情報と紐づけることは容易ではない。しかしながら、実務的には「何故特定のイベントに関するこの企業のスコアが高いのか」という問いに答える必要のある局面は多く、予測のための判別器が完全にブラックボックスになっている状況は実務的に望ましいものとは言えない。

そこで、本稿では、variable importance (Janitza et al. 2018) を参照することで、特定の変数が検知・予測における分類精度の向上への程度貢献しているかを参照することで、こうした批判へ部分的に対応する。更に、サンプル企業の平均的な属性からスタートして各変数を僅かに変化させた場合において、対応した S がどの様に変化するかを図示することで、 X に含まれる特定の情報の変動がもたらす予測上の定性的な含意を確認する。

注意すべきは、こうした試みが、他の要因を固定した場合に、特定の要因が不正会計の発生をもたらす「限界的な効果」の識別を必ずしも目的とはしていないという点である。例えば、 X に含まれる特定の変数が高い variable importance を有しており、かつ当該変数の上昇が S の上昇に繋がっていることを確認したとしよう。検知・予測の文脈でこの結果が意味するところは、当該変数が上昇した場合において、現時点（検知）もしくは将来時点（予測）において不正会計の発生がより高い確率で観察されるというものである。しかし、この結果は、当該変数と相関の高い何らかの変数が、不正会計イベントと強く相関していることから生み出された結果である可能性も排除できない。こうした spurious correlation の可能性を認識した上で、結果を慎重に解釈する必要がある。

⁵ 例えば、recall 指標は、特定の閾値以上のスコアが与えられたサンプルについてイベントが現在生じている（将来的に生じる）と検知（予測）するという設定の下で、「実際にイベントが現在生じている（将来的に生じる）サンプルのうち、どの程度の割合を正しく検知（予測）出来たか」という指標である。この指標は、実際のイベント発生を可能な限り見逃さないことを重視した指標と言え、例えば、重大な疾病を検知する際には有用である。しかし、上記の閾値を可能な限り低く設定すればこの recall 指標は必ず 100% となるため、特定の指標のみに依拠する形で精度評価を行う事には問題がある。

⁶ ROC 曲線と AUC に関しては、補論 1 参照。

Xに含まれる特定の変数がSに与える causal impact の推定については、近年様々な手法が提案されており (Chernozhukov et al. 2018)、応用事例も徐々に蓄積されている (Miyakawa 2019b)。こうした因果関係の識別を正確に行うことが出来れば、予測に比してより pro active な取り組みが可能となる。例えば、特定のガバナンス構造を採用したり、規制を導入したり、などの能動的な働きかけが不正会計の発生を防ぐことが確認されれば、実務的・政策的な観点から重要な情報となるだろう。この点については今後の検討課題としたい。

5. データ

本節では、分析用データセットについて概観する。本稿の分析では、本邦証券市場に上場し、会計基準に日本基準を適用している企業（金融業を除く）に限定した分析を行う。

5.1 不正会計フラグ

本稿の分析では、上場企業の有価証券報告書、金融庁による有価証券報告書等の虚偽記載に対する課徴金納付命令の有無、および証券取引等監視委員会が行った虚偽有価証券報告書提出事件としての検察庁への告発の有無を利用する。これらのデータを参照することで、上場企業の財務諸表について、過去にどのような訂正が行われたかを計測することが可能となり、本邦企業による財務情報に関するミスレポーターリングを網羅的に計測することができる。また、これらのミスレポーターリングが、金融庁から有価証券報告書等の虚偽記載に対する課徴金納付命令が出された会計年度において行われていたものか否か、また、証券取引等監視委員会から虚偽有価証券報告書提出事件として告発があった会計年度において行われていたものか否か、を判別することができる点が挙げられる。こうした情報を用いることで、公的な制裁が発動されるような重大な不正会計イベントが生じていた企業×年度のペアを特定することが可能となる。分析に当たっては、この制裁情報を用いて企業×年度のペアに対して設定した下記の不正会計フラグ1に加えて、こうした制裁の対象になっていない訂正レコードを追加的に参照して計測した不正フラグ2を用いる。

不正フラグ2を対象とした検知・予測については、不正会計としての重要度が相対的に低いイベントを対象としているという意味で相対的に困難なタスクといえる。本稿において上記の二種類のフラグを用いた分析を行う狙いの一つとして、こうした検知・予測の意味で困難なタスクにおいてどの程度の精度を実現することができるかを検討する点にある。

不正会計フラグ1：

以下の2条件 (①、②) のいずれかを満たす企業×年度に1 (ポジティブデータ)、それ以外の場合に0 (ネガティブデータ) とするダミー変数

①金融庁から有価証券報告書等の虚偽記載に対する課徴金納付命令が出された会計年度のうち、決算書の訂正がある会計年度

②証券取引等監視委員会から有価証券報告書等の虚偽記載に対する告発があった会計年度のうち、決算書の訂正がある会計年度

不正会計フラグ2：

不正フラグ1の2条件(①、②)に以下の条件③を加えた3条件(①、②、③)のいずれかを満たす企業×年度に1(ポジティブデータ)、それ以外の場合に0(ネガティブデータ)とするダミー変数

③金融庁または証券取引監視委員会から課徴金納付命令または告発を受けた企業のうち、命令・告発を受けた会計年度以外の会計年度において、決算書の訂正があった会計年度

上記の定義から明らかなように、不正会計フラグ1は重大な不正会計イベントのみを計測対象としたものである一方、不正会計フラグ2は当該不正会計イベントに付随して修正が必要となったミスレポートを含むものである。実務的には、不正会計フラグ1に対応するイベントを検知・予測することが最重要課題と考えられるが、仮に課徴金納付命令や虚偽記載に関する告発が無かった年度であっても、そうした重要な不正会計イベントに付随するミスレポートを検知・予測することが出来れば、経営管理や監査実務の面では重要な情報となり得るだろう。

図1はこれらのフラグによる不正会計の観測数(縦軸)を各年毎(横軸)に示したものである。これを見ると分析対象期間の後半でフラグ数が減少していることが分かる。これは、一般的に既述の制度に基づく不正会計の発覚には、実際の不正タイミングから一定の時間を要するため、そうした発覚までの時間が確保できていないことによると考えられる。このことは、分析期間の前半のデータを用いて行うトレーニングステップに比べて、後半の期間のデータを用いて行うテストステップにおいて、不正会計の観測数(ポジティブデータ数)が相対的に少ないことを意味している。こうしたトレーニングステップとテストステップの間でのフラグの質的な差異が結果に致命的な影響をもたらしていないことを確認する趣旨から、本稿では、トレーニングステップとテストステップを同期間における異なる企業群に基づいて行う頑健性のチェックも行う。

5.2 変数

本稿の分析において検知・予測を行うために参照する変数は、以下の三つのデータソースから構築される。第一に、5.1節で説明したデータから各上場企業の各年度における財務諸表に格納されている財務情報を抽出し変数を構築する。変数の構築に当たっては、まず、本邦企業の不正会計イベント検知に関する代表的な先行研究である Song et al. (2016)で用いられている変数をまず準備する。具体的には、無形資産から現預金を控除した値を総資産で除したソフト資産比率、裁量の会計発生高、株式及び債券の発行ダミー変数、異常リターンに関

するスコア、実体的裁量行動、非金融法人による持ち株比率、である。これらの変数は、Dechow et al. (2011)などで議論されている、利益の質、パフォーマンス、非金融指標、オフバランス活動、市場関連変数といった、理論的に見て不正会計の発生と関連している可能性の高い変数へ、実物経済活動に関するマニピレーションの度合い、会計上の保守主義、日本特有の要因を追加したものである。なお、これらの変数が何らかの事情により欠損している場合については、当該欠損値を0で補完した上で、別途各変数の欠損に対応するダミー変数を設定している⁷。

これらの変数が会計領域の理論的な検討を踏まえた尤もらしい変数である一方、本稿の目的である検知・予測の観点からは、より高次元の企業情報を変数として参照することでより高い検知・予測精度を実現できる可能性もある。本稿では、既述の財務諸表データを用いて、40種類程度の財務変数と対応する欠損ダミー変数を追加的に参照する。売上・売掛債権、利益、キャッシュフロー、たな卸資産、固定資産、繰延税金資産・負債、総資産・純資産、その他の変数で、具体的には、売上・売掛債権に関してそれぞれの変化や債権回転期間、海外売上比率など、利益に関して各段階損益のレベル・変化・利益率など、キャッシュフローに関してそのレベル、営業利益との関係など、たな卸資産に関して回転期間や純資産・総資産比率、固定資産に関して有形固定資産のレベル・総資産比率、のれん純資産比率など、繰延税金資産・負債に関して利益剰余金との関係、総資産・純資産に関してレベルや変化、その他に関して新規上場か否かに関する変数である。

更に、宋(2018)などが指摘するガバナンス変数の潜在的な重要性を勘案して、日経 NEEDs Financial Quest データから、以下の二種類の変数群を構築した。第一に、外国法人、役員、大株主の持ち株比率、従業員の平均勤続変数からなるガバナンス変数である。第二に、各企業の金融機関取引関係から構築した、取引銀行数、メガバンクの借入比率、借入シェアのハーフィンダール指数に加えて、それらの変化を用いた。表1はこれらの変数群を要約したものである。不正会計の検知・予測モデルの構築に当たっては、これらの変数に加えて、32業種の産業分類を示す業種ダミーを加えている。表2は、32業種の産業分類を示したものである。

こうしたデータ構築の結果、例えば、不正会計フラグ1を対象とした検知モデルの構築に当たって、2006年1月決算期から2016年3月期決算までの4,094社に関する34,923観測値(企業×年度)を用いる。なお、このサンプル内で不正会計フラグ1がポジティブとなっている観測値は126観測値である。このケースにおいては、モデルのトレーニングには2006年1月期～2013年3月期における25,401観測値を用い、モデルのテストには2013年4月期～2016年3月期における9,522観測値を用いる。不正会計フラグ1のポジティブデータは各々107観測値と19観測値である。表3は、このケースを含む全体の観測値を要約した

⁷ こうした欠損値の補完については、この他にも、線型補完や観測されているサンプルの平均値・中位値で補完するなどの方法が考えられる。実際にサンプルの平均値で補完したデータセットを用いて、本稿で実施した分析を行ったところ同様の結果が得られた。分析結果については、補論2を参照。

ものである。

6. 予測分析

本節では、不正会計の検知・予測モデルの構築、精度評価、結果の解釈について述べる。

6.1 トレーニングステップ

第一に、不正会計の検知を目的としたモデルのトレーニングを行う。不正会計の計測が行われた年度と同時点において計測済みの企業情報 X を参照した上で、**Weighted Random Forest**の枠組みでモデルのトレーニングを行うが、その際、投入する変数群の種別によってモデル1からモデル16までを設定する。

表4はこの投入変数群に関する詳細をまとめたものであるが、特に注目すべきモデルとしてモデル1、2、6、12を取り上げる。モデル1は、本稿における比較対象であり、不正会計の検知を目的とした代表的な先行研究である **Song et al. (2016)**のセットアップに基づき、理論に基づく限定的な変数をパラメトリックなモデルを用いることで不正会計の検知モデルを構築したものである。これに対して、モデル2は、モデル1と同様の変数群を参照した上で、モデルを **Weighted Random Forest** に変更したものであり、限定的な変数セットアップの下で機械学習手法が実現する検知精度の改善度合いを確認するものである。モデル6は、こうした機械学習ベースのモデル構築によって可能となる高次元の変数利用を行ったものであり、既存研究の変数と追加的な財務変数の両方を用いている。当該モデル6の検知精度をモデル2と比較することで、追加的な変数の利用がもたらす効果を把握することができる。最後に、モデル12は、既述の追加的な財務変数に加えてガバナンス変数及び銀行取引変数を含めたものである。

第二に、同様のセットアップで不正会計の一期先予測を行う。検知モデルのトレーニングとの相違点は、不正会計の計測が行われた年度の前年度時点において計測済みの企業情報 X を参照する点にある。これらの不正会計検知・予測モデルのトレーニングに際して、不正会計フラグ1と2の各々を用いる。

6.2 テストステップ

表5は、不正会計の検知と予測について、不正会計フラグ1と2を用いてトレーニングしたモデルについて、テストデータにおける精度評価指標である **AUC** をその標準誤差と併せて表記したものである。図2の上段及び下段のパネルは、重大+付随イベントを対象とする検知と予測について、各モデルの **AUC** をその95%信頼区間と併せて表記したものである。

第一に、検知と予測の両方において、モデル1からモデル2にかけて精度の若干の上昇が確認される。このことは投入する情報が限定的な量に留まっている場合でも、複雑な非線形性を許容するノンパラメトリックモデルを用いることで、**hold-out data** での検知・予測精度が

向上することを示唆している。なお、AUC の計測値に関してこうした上昇がみられる一方で、その標準誤差を勘案すると、両モデルの AUC に差がないという帰無仮説を棄却することは出来ない点に注意が必要である。

第二に、モデル 2 からモデル 6 にかけて大幅な AUC の上昇が確認されている。この結果は、予測モデルの推定手法を変更したことによって可能となった高次元情報（追加的な変数）の利用が、検知・予測精度の改善に当たって重要な役割を示していることを意味している。

第三に、こうした結果の一方で、モデル 6（既存研究の変数および追加的な財務変数）からモデル 1 2（既存研究の変数、追加的な財務変数、ガバナンス変数および銀行変数）にかけては、検知・予測精度の上昇が限定的な水準にとどまっていることも確認できる。この結果については、既存研究の変数とガバナンス変数に基づいたモデル 7 の精度や、既存研究の変数と銀行変数に基づいたモデル 8 の精度が、モデル 2 に比して改善していることと併せて解釈する必要がある。一つの解釈として、ガバナンス変数および銀行変数が持つ情報の大半は、追加的な財務変数に包含されている（追加的な財務変数が持つ情報と強く関連している）ものの、不正と関連する有用な情報を含んでいると考えることができる。

第四に、表 5 の結果から、検知及び予測の難易度が相対的に高いと考えられる不正フラグ 2 について、従来モデルに比べて予測精度が顕著に向上している。この結果は、検知・予測が困難な対象について、高次元の変数を用いることの意義が高いことを示唆している。

図 3 は、モデル 1 2 に基づいて計測されたスコアと実際の会計不正の発生確率の対応関係を図示したものである。作図に当たっては、実際の不正会計フラグを **Weighted Random Forest** のスコアへ回帰することで対応関係を推定し、その結果を、実際の不正発生確率と併せて図示している。第一に、スコアベースの不正確率と実際の不正確率とが概ね同様の動きを示しており、機械学習ベースのモデルから出力されたスコアの経済的な意味合いを解釈することが可能であることが分かる。第二に、特にスコアが高いグループにおいて不正の発生確率が大きくに上昇していることが分かる。

最後に、本節で示した結果の頑健性を確認するために、トレーニングデータとテストデータを時間方向で分割する現在のサンプル分割方法とは異なる方法でトレーニングデータとテストデータを設定した上でモデル構築と精度評価を行った結果を示す。こうした分析を行う理由は、これまでの分析が複数の期間に亘って不正会計フラグがポジティブとなっている企業（特に不正会計フラグ 2）が存在するため、企業に関して時不変の要因が存在し、かつトレーニングデータとテストデータの両方に当該企業が含まれる場合に、モデルの検知・予測精度の計測に誤りが生じる可能性があるためである。図 4 は、各企業がトレーニングデータもしくはテストデータの何れかのみ含まれるように分割したデータを複数回（100 回）設定し、AUC の経験分布を示したものである。これまでの結果と同様の含意が得られていることが確認できる。

6.3 結果の解釈

前節の結果から、機械学習ベースの検知・予測モデルにおいて特に財務変数とガバナンス変数を投入したモデルのトレーニングを行うことで、**hold-out data** で十分な精度を実現できることを確認した。本節では、この学習済み検知・予測モデルの含意を、**variable importance** と変数の **perturbation** によって確認する。

まず、表6は、モデル2、6、12に関して、**variable importance** を表記したものである。第一に、追加的な財務変数の検知・予測に対する貢献が相対的に高いことが確認される。第二に、既存研究では十分に勘案されていなかったガバナンス変数の一部が **variable importance** の意味で上位にランクされていることが分かる。

次に、図5は、平均的な企業属性を始点として、各属性を上下に変化させた際のスコアの変化を図示したものである。第一にソフト資産比率に代表される既存研究が重視してきた変数が、既存研究の実証結果と整合的な含意を有していることが確認できる。第二に、平均勤続年数に代表されるガバナンス関連変数やたな卸資産純資産比率に代表される追加的な財務変数が、直観と整合的な方向でスコアに影響していることが分かる。第三に、売上高・営業収益に代表されるように、スコアへの影響が単調ではないものも確認されており、複雑な非線形性を考慮できる機械学習手法の利用価値が高いことが示唆されている。

7. ディスカッション

本節では、前節までに確認した本稿での実証結果を踏まえたディスカッションを行う。第一に、本稿で用いた機械学習手法ベースの検知・予測モデル構築については、特にトレーニングステップにおいて様々な **hyper parameter** の設定が必要であることに注意が必要となる。例えば、構築するツリーの数、分岐数の最小値、分岐ルールの設定に当たって参照する統計量の種別などである。補論3ではこれらに関する頑健性を確認しており、本稿で得られた結果がこれらのパラメータの設定に大きく依存しないことを確認しているが、機械学習手法を用いることで予測モデル構築にかかる全ての作業を自動化できるという誤った理解は避ける必要がある。

第二に、本稿で採用した機械学習手法以外にも、不正会計の検知・予測に利用可能な手法が多く存在する点は認識しておく必要がある。実際に、本稿の執筆メンバーの一部が異なる機械学習手法を用いて行った分析においても、良好な検知・予測精度が実現されており、将来的には異なる手法から得られた検知・予測スコアを再度合算することでより頑健性の高い検知・予測を行うことも一案と考えられる。逆に、本稿で得られた知見を活用することで、検知・予測の観点から相対的に重要と考えられる限られた数の変数群を特定し、使い勝手の良いコンパクトな検知・予測モデルを構築することも有益であろう。実際に、補論4において示している通り、分析対象のフラグが共通している場合でも、検知と予測において高い **variable importance** を示す変数群には差異が存在するほか、予測という共通した分析を行う

場合でもフラグの設定によって高い **variable importance** を示す変数群には差異が生じる。こうした観察事実は、上記のコンパクトなモデル構築において目的に応じた変数選択がなされるべきであることを意味している。

第三に、本稿で構築した検知・予測モデルは社会・経済環境、会計基準が不変であることを前提にしている。検知・予測モデルを運用するにあたっては、これらの変化がモデルに及ぼす影響を常に意識する必要がある。例えば、新たなビジネスにおいてこれまでと全く異なる不正が行われた場合に、現状のモデルではこれを検知・予測できない可能性がある。また、企業の実態が同一であっても会計基準の変化により財務変数が異なる値を取り、不正検知・予測結果に影響を及ぼす可能性がある。このような問題に対応するには、モデルの有効性を定期的にモニタリングする仕組みを構築し、継続的にモニタリングを実施することが重要である。

第四に、高精度の予測スコアが構築できた場合、当該スコアをどの様に活用するかが次なる重要な問題となる点を十分に認識しておく必要がある。高い確率で不正会計が発生している、もしくは発生する可能性があると考えられた企業が存在するとして、実際にどのような点において不正会計の蓋然性が高く、どのような対応を現時点で必要とするかを検討するためには、現時点では専門的な知見を持ったスタッフによる高度な判断が必要となる。この点に関連して、本稿で実施した企業レベルの不正会計検知・予測を補完する形で、勘定科目レベルの異常検知が重要な研究テーマと考えられる。また、実際に検知・予測スコアが提供された場合に現場のスタッフがどの程度そうしたスコアに依拠した業務運営を行うかという問題が存在する点も重要となる（例：Kleinberg et al. 2018）。

第五に、政策的な観点からこうした検知・予測技術の発展をどの様に捉えるべきかという問題が存在する。例えば、不正会計の検知を対象とするモデルの詳細が完全に公知となった場合、何らかの事情から不正会計を行おうとする企業が、そうしたモデルに検知されにくい形での不正を試みる可能性もあるだろう。現実には、モデルの詳細が完全に共有されることは考えにくく、またモデルに検知されない形での不正の実施が必ずしも容易ではない可能性も高いため、こうしたある種の「イタチごっこ」がどの程度現実的な問題となるかは不明であるが、少なくとも政策的見地からは認識しておくべき論点と言える。例えば、ある単階で確立した検知・予測モデルを無効化するような企業サイドの悪意ある取り組みを受けたモデルの不断の改善に取り組むというのが一つのスタンスと考えられる。

第六に、本稿では十分に取り扱わなかった因果推論の方向で議論を深める必要がある。特定の監査上の取り組みやガバナンス上の工夫を採用することで、因果関係として不正イベントの発生を抑えることができるか否かは、不正会計を未然に防ぐための方策を検討する意味で重要である。

8. まとめ

本研究は、不正会計の検知と予測における機械学習手法の有用性を実証的に検討したものである。第一に、既存研究における限定的な変数に基づくパラメトリックモデルに代えて機械学習手法を用いることで、高次元の変数に基づく不正会計の「検知」を行い、その精度の改善を検証する。第二に、既存研究において十分に検討されていない不正会計の「予測」を同様のフレームワークで行う。第三に、構築された予測モデルの解釈を行う目的から、不正会計の予測結果（スコア）が、どのような情報に反応しているかを検討する。本邦上場企業の不正会計イベントを対象とした実証分析から、機械学習手法の利用が検知精度の改善に一部貢献するほか、当該手法の利用によって参照可能となる高次元の変数利用が検知精度の大幅な改善に貢献していることを確認した。また、機械学習ベースのモデルが実務的観点から十分な水準の予測精度を実現できることを確認した。なお、本研究では、既存研究が理論的な想定に基づいて参照してきた変数以外の変数（例：ガバナンス関連変数）について、不正会計の検知・予測にあたって一定程度有用な情報が含まれていることも確認した。この結果は、不正会計の発生メカニズムに関する今後の理論的検討に対しても有益な情報を提供している。

参考文献

- 宋 (2018)「昨今の経済環境等の変化に対応した不適正会計の早期発見に関する調査・研究」
FSA Institute Discussion Paper Series, DP 2017-6.
- Breiman, Leo. 2001. Random Forests. *Machine Learning* 45 (1): 5–32.
- Chen, C., A. Liaw, and L. Breiman. 2004. Using Random Forest to Learn Imbalanced Data. Technical Report 666 Statistics Department of University of California at Berkeley.
- Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins. 2018. Double/Debiased Machine Learning for Treatment and Structural Parameters. *Econometrics Journal* 21(1): 1-68.
- Dechow, P. M., Ge, W., Larson, C. R. and R.G. Sloan. 2011. Predicting Material Accounting Misstatements. *Contemporary Accounting Research* 28(1): 17-82.
- Dechow, P. M., Ge, W., and C. Schrand. 2010. Understanding Earnings Quality: A Review of the Proxies, their Determinants and their Consequences. *Journal of Accounting and Economics* 50 (2)(3): 344-401.
- Dechow, P. M., R. G. Sloan, and A. P. Sweeney. 1996. Causes and Consequences of Earnings Misstatement: An Analysis of Firms Subject to Enforcement Actions by the SEC. *Contemporary Accounting Research* 13(1): 1-36.
- Janitza, S., E. Celik, Anne-Laure Boulesteix. 2018. “A Computationally Fast Variable Importance Test for Random Forests for High-Dimensional Data,” *Advances in Data Analysis and Classification* 12(4): 885-915.
- Perols, J. 2011. “Financial Statement Fraud Detection: An Analysis of Statistical and Machine Learning Algorithms,” *Auditing: A Journal of Practice & Theory* 30(2): 19-50.
- Perols, J., Bowen, B., and C. Zimmerman, 2017. “Finding Needles in a Haystack: Using Data Analytics to Improve Fraud Prediction,” *Accounting Review* 92(2): 221-245.
- Song, M., N. Oshiro, and A. Shuto. 2016. Predicting Accounting Fraud: Evidence from Japan. *The Japanese Accounting Review* 6: pp.17-63.
- Kleinberg, J., H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan. 2018. Human Decisions and Machine Predictions. *Quarterly Journal of Economics* 133(1): 237-293.
- Miyakawa, D. 2019a. “Forecasting Firm Dynamics with Machine Learning,” mimeo.
- Miyakawa, D. 2019b. “Shocks to Supply Chain Network and Firm Dynamics: An Application of Double Machine Learning,” forthcoming in RIETI Discussion Paper Series.
- West, J. and M. Bhattacharya. 2016. “Intelligent Financial Fraud Detection: A Comprehensive Review,” *Computer & Security* 57: 47-66.

図表

表 1 変数定義と変数区分

変数大区分	変数小区分	変数名
Basic		ソフト資産比率
		裁量的会計発生高
		株式及び債券を発行した場合に1、それ以外は0を取るダミー変数
		異常リターンに関するスコア
		裁量的営業キャッシュフロー
		非金融法人による持ち株比率
Additional	売上・売掛債権	売上高
		売上高の前期比
		売上高の3期比較
		売上高の増加率
		海外売上高比率
		売掛債権残高の前期との差
		売掛債権残高が前期より増加した場合に1、それ以外は0とするダミー変数
		売掛債権の回転期間
		売掛債権の前期回転期間
		売掛債権の回転期間前期比
		売掛債権の増減率
		売掛債権の増減率を売上高の増加率で除した比率
		利益
営業利益率		
営業利益が2期連続赤字の場合に1、それ以外は0とするダミー変数		
経常利益		
経常利益率		
営業外損益		
営業外損益率		
親会社株主に帰属する当期純利益／当期利益の3期比較		
経常損失で当期純利益または経常利益で当期純損失の場合に1、それ以外は0とするダミー変数		
キャッシュ・フロー		営業キャッシュ・フロー
		2期連続営業キャッシュ・フローがマイナスの場合に1、それ以外は0とするダミー変数
		営業利益対営業キャッシュ・フロー比
たな卸資産		たな卸資産回転率
		たな卸資産純資産比率
		たな卸資産総資産比率
固定資産		有形固定資産残高
		有形固定資産総資産比率
		のれん純資産比率
		のれん残高＞営業利益の場合に1、それ以外は0とするダミー変数
繰延税金資産・負債		利益剰余金がマイナスで繰延税金資産が計上されている場合に1、それ以外は0とするダミー変数
		繰延税金資産・負債(純額)利益剰余金比率
総資産・純資産		純資産の3期比較
		総資産残高
その他		新規上場の場合に1、それ以外は0とするダミー変数

表 1 (続き) 変数定義と変数区分

変数大区分	変数小区分	変数名
Governance		外国法人等所有株式数の割合
		役員所有株式数の割合
		従業員平均勤続年数
		大株主所有株式数の割合
Bank Relation		長期借入金の取引銀行数
		短期借入金の取引銀行数
		借入金合計の取引銀行数
		長期借入金のメガバンク借入比率
		短期借入金のメガバンク借入比率
		借入金合計のメガバンク借入比率
		長期借入金シェアのハーフィンダール指数
		短期借入金シェアのハーフィンダール指数
		借入金合計シェアのハーフィンダール指数
		企業規模調整済み長期借入金シェアのハーフィンダール指数
		企業規模調整済み短期借入金シェアのハーフィンダール指数
		企業規模調整済み借入金合計シェアのハーフィンダール指数
		長期借入金を1行のみから調達している場合に1、それ以外は0とするダミー変数
		短期借入金を1行のみから調達している場合に1、それ以外は0とするダミー変数
		借入金合計を1行のみから調達している場合に1、それ以外は0とするダミー変数
		長期借入金の取引銀行数の前期との差
		短期借入金の取引銀行数の前期との差
		借入金合計の取引銀行数の前期との差
		長期借入金のメガバンク借入比率の前期との差
		短期借入金のメガバンク借入比率の前期との差
		借入金合計のメガバンク借入比率の前期との差
		長期借入金シェアのハーフィンダール指数の前期との差
		短期借入金シェアのハーフィンダール指数の前期との差
		借入金合計シェアのハーフィンダール指数の前期との差
		企業規模調整済み長期借入金シェアのハーフィンダール指数の前期との差
		企業規模調整済み短期借入金シェアのハーフィンダール指数の前期との差
		企業規模調整済み借入金合計シェアのハーフィンダール指数の前期との差
		長期借入シェア最上位の銀行が交代した場合に1、それ以外は0とするダミー変数
		短期借入シェア最上位の銀行が交代した場合に1、それ以外は0とするダミー変数
		借入金合計シェア最上位の銀行が交代した場合に1、それ以外は0とするダミー変数

表2 32業種の産業分類

No.	業種名	No.	業種名	No.	業種名	No.	業種名
1	食品	9	鉄鋼	17	その他製造	25	陸運
2	繊維	10	非鉄金属製品	18	水産	26	海運
3	パルプ・紙	11	機械	19	鉱業	27	空運
4	化学	12	電気機器	20	建設	28	倉庫
5	医薬品	13	造船	21	商社	29	通信
6	石油	14	自動車	22	小売業	30	電力
7	ゴム	15	輸送用機器	23	不動産	31	ガス
8	窯業	16	精密機器	24	鉄道・バス	32	サービス

表3 観測値数一覧

		検知			予測		
		トレーニング	テスト	全データ	トレーニング	テスト	全データ
不正会計 フラグ1	財務諸表観測数	25,401	9,522	34,923	25,401	9,522	34,923
	会社数	3,858	3,424	4,094	3,858	3,424	4,094
	ポジティブ観測数	107	19	126	107	19	126
不正会計 フラグ2	財務諸表観測数	25,401	9,522	34,923	25,401	9,522	34,923
	会社数	3,858	3,424	4,094	3,858	3,424	4,094
	ポジティブ観測数	150	23	173	150	23	173

表4 モデル一覧

変数グループ	定義	モデル															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
		Probit	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF
Basic	6変数(既存研究による)	○	○				○	○	○	○	○	○					
Additional	144変数			○			○			○	○	○	○	○	○	○	
Governance	16変数				○			○		○		○	○	○		○	○
Bank relation	30変数					○			○		○	○	○		○	○	○
industry dummy	32変数	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

表5 精度評価結果

結果パターン	モデル															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	Probit	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF	WRF
検知 + 不正会計フラグ1																
AUC	0.71	0.79	0.84	0.64	0.68	0.86	0.81	0.86	0.86	0.88	0.86	0.87	0.82	0.86	0.85	0.74
s.e.	0.06	0.05	0.04	0.07	0.06	0.04	0.05	0.03	0.04	0.04	0.03	0.04	0.05	0.04	0.04	0.06
検知 + 不正会計フラグ2																
AUC	0.73	0.77	0.86	0.67	0.75	0.88	0.78	0.82	0.88	0.89	0.84	0.90	0.86	0.88	0.88	0.78
s.e.	0.04	0.05	0.03	0.06	0.05	0.03	0.05	0.04	0.04	0.04	0.04	0.04	0.03	0.04	0.04	0.05
予測 + 不正会計フラグ1																
AUC	0.68	0.68	0.80	0.64	0.60	0.80	0.72	0.73	0.80	0.83	0.75	0.82	0.79	0.84	0.82	0.67
s.e.	0.07	0.06	0.05	0.07	0.07	0.05	0.06	0.05	0.06	0.04	0.05	0.05	0.06	0.05	0.05	0.07
予測 + 不正会計フラグ2																
AUC	0.65	0.68	0.83	0.70	0.67	0.82	0.73	0.75	0.82	0.85	0.76	0.85	0.82	0.85	0.85	0.72
s.e.	0.06	0.05	0.04	0.05	0.06	0.04	0.04	0.05	0.04	0.04	0.05	0.04	0.04	0.04	0.04	0.05

備考: 太字はAUC0.9以上

表6 Variable importance

モデル2 Basic + WRF

Rank	変数区分	変数名	重要度
1	Basic	非金融法人による持ち株比率	972.3
2	Basic	ソフト資産比率	946.1
3	Basic	株式及び債券を発行した場合に1、それ以外は0を取るダミー変数	790.7
4	Basic	非裁量的営業キャッシュフロー	636.2
5	Basic	異常リターンに関するスコア	608.5
6	Basic	非裁量的会計発生高	534.5

モデル6 Basic & Add + WRF

Rank	Rank (モデル2)	変数区分	変数名	重要度
1	-	Additional	営業外損益率	363.7
2	-	Additional	有形固定資産総資産比率	334.4
3	-	Additional	有形固定資産残高	333.0
4	3	Basic	株式及び債券を発行した場合に1、それ以外は0を取るダミー変数	298.4
5	-	Additional	営業外損益	290.1
6	-	Additional	売上高	288.6
7	-	Additional	たな卸資産純資産比率	286.4
8	2	Basic	ソフト資産比率	279.3
9	1	Basic	非金融法人による持ち株比率	264.5
10	-	Additional	営業利益対営業キャッシュ・フロー比	260.5
11	-	Additional	たな卸資産回転率	212.4
12	-	Additional	営業キャッシュ・フロー	208.4
13	-	Additional	総資産残高	207.2
14	-	Additional	たな卸資産総資産比率	201.6
15	4	Basic	非裁量的営業キャッシュフロー	190.5
16	-	Additional	経常利益率	175.2
17	-	Additional	繰延税金資産・負債(純額)利益剰余金比率	171.5
18	-	Additional	経常利益	161.3
19	-	Additional	売掛債権の回転期間	147.2
20	-	Additional	営業利益率	142.6
21	-	Additional	のれん純資産比率	136.9
22	-	Additional	営業利益	125.6
23	5	Basic	異常リターンに関するスコア	120.6
24	-	Additional	売掛債権の前期回転期間	115.7
25	6	Basic	非裁量的会計発生高	108.4

モデル12 Basic, Add, Gov & BK + WRF

Rank	Rank (モデル2)	Rank (モデル6)	変数区分	変数名	重要度
1	-	1	Additional	営業外損益率	267.3
2	-	3	Additional	有形固定資産残高	265.7
3	-	2	Additional	有形固定資産総資産比率	257.4
4	3	4	Basic	株式及び債券を発行した場合に1、それ以外は0を取るダミー変数	216.1
5	-	-	Governance	従業員平均勤続年数	213.8
6	-	6	Additional	売上高	213.4
7	-	5	Additional	営業外損益	209.5
8	1	9	Basic	非金融法人による持ち株比率	207.2
9	2	8	Basic	ソフト資産比率	204.6
10	-	7	Additional	たな卸資産純資産比率	196.9
11	-	10	Additional	営業利益対営業キャッシュ・フロー比	186.3
12	-	-	Governance	役員所有株式数の割合	169.0
13	-	12	Additional	営業キャッシュ・フロー	167.6
14	-	13	Additional	総資産残高	154.7
15	-	-	Governance	大株主所有株式数の割合	153.4
16	4	15	Basic	非裁量的営業キャッシュフロー	149.0
17	-	11	Additional	たな卸資産回転率	145.0
18	-	14	Additional	たな卸資産総資産比率	144.9
19	-	17	Additional	繰延税金資産・負債(純額)利益剰余金比率	126.4
20	-	16	Additional	経常利益率	126.3
21	-	18	Additional	経常利益	120.6
22	-	22	Additional	営業利益	108.1
23	-	19	Additional	売掛債権の回転期間	105.3
24	-	20	Additional	営業利益率	103.0
25	-	21	Additional	のれん純資産比率	93.4

図 1

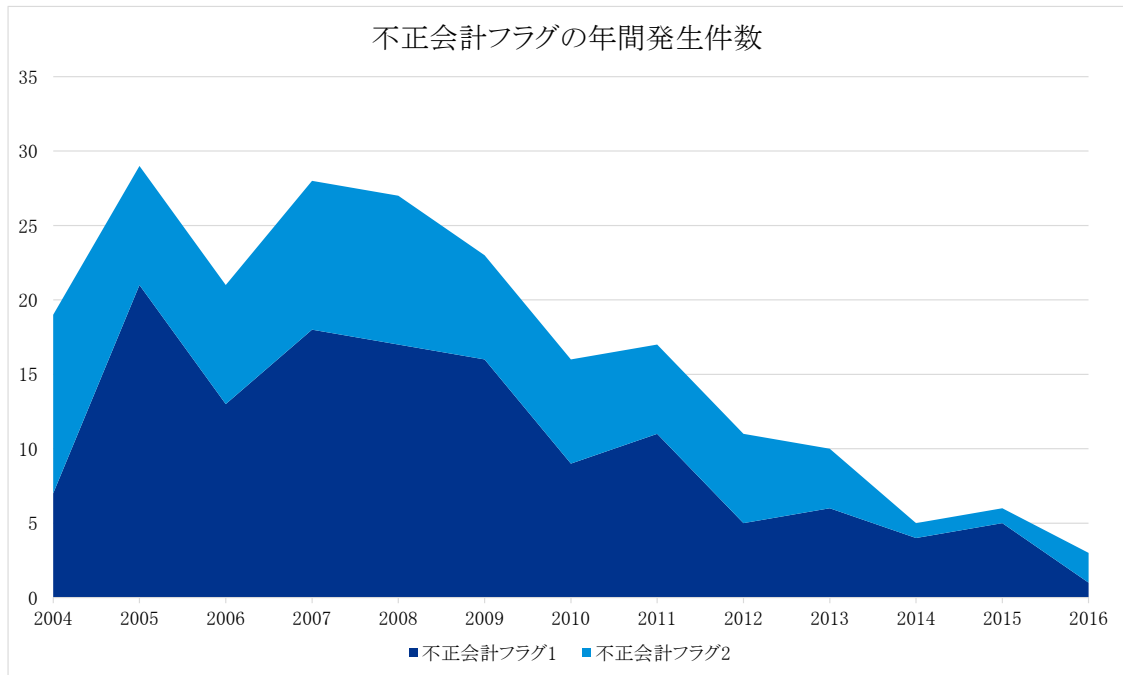


図2 精度評価結果

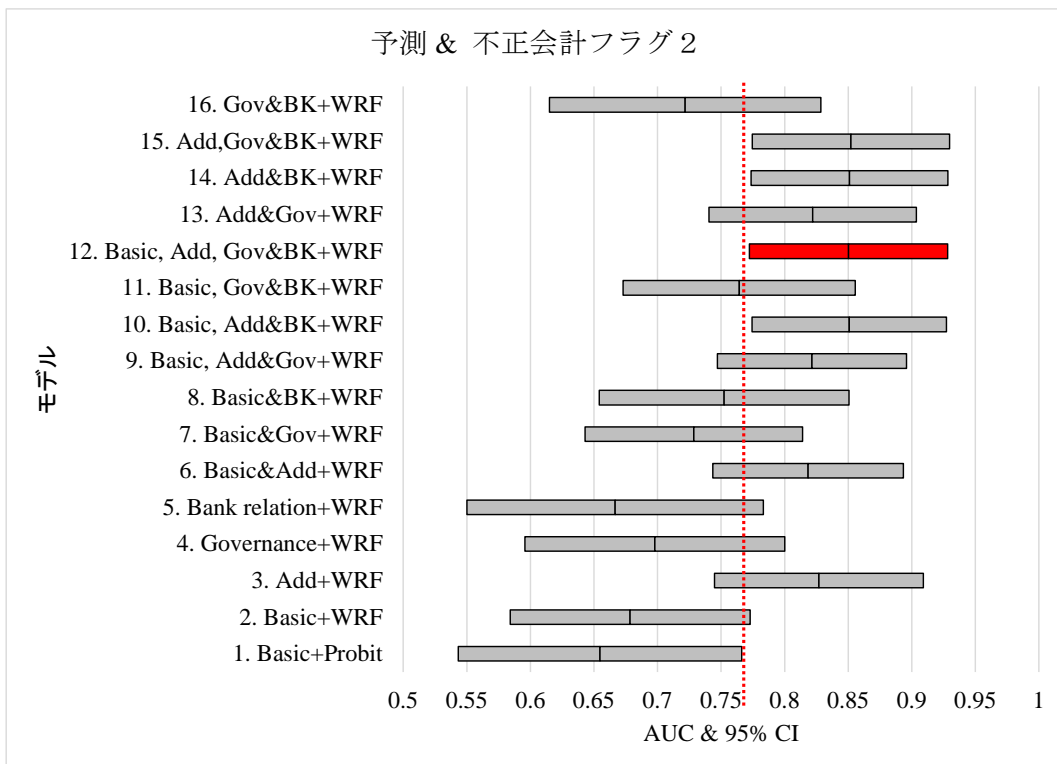
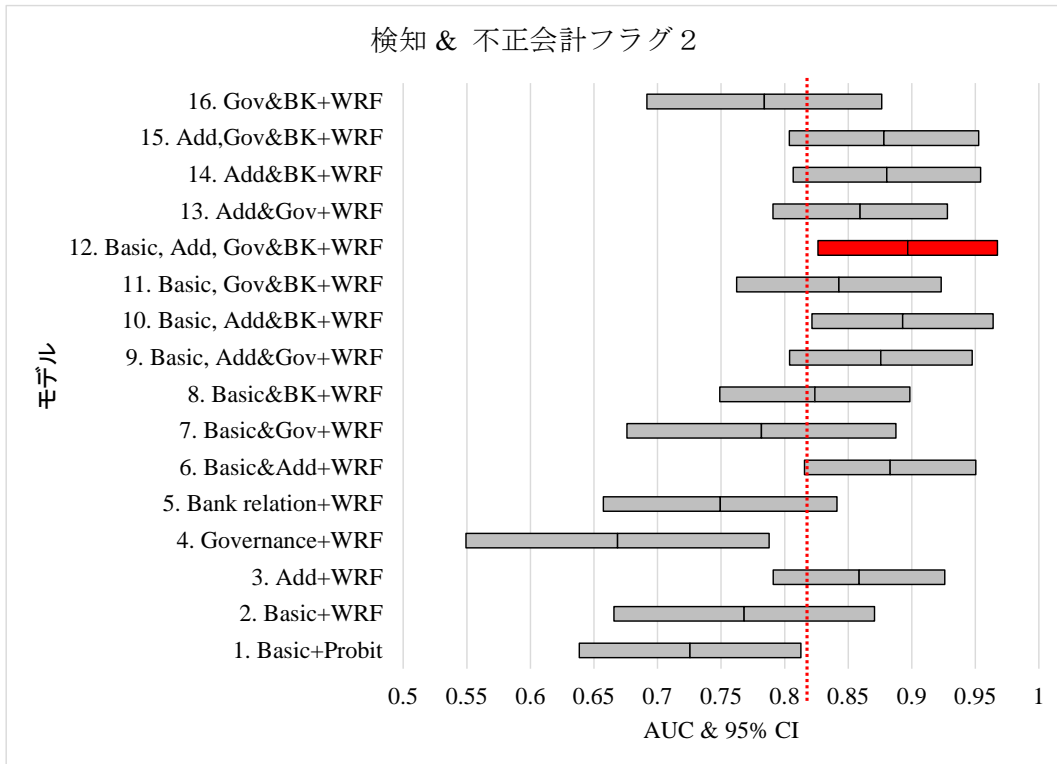


図3 スコアと不正会計発生確率の対応関係

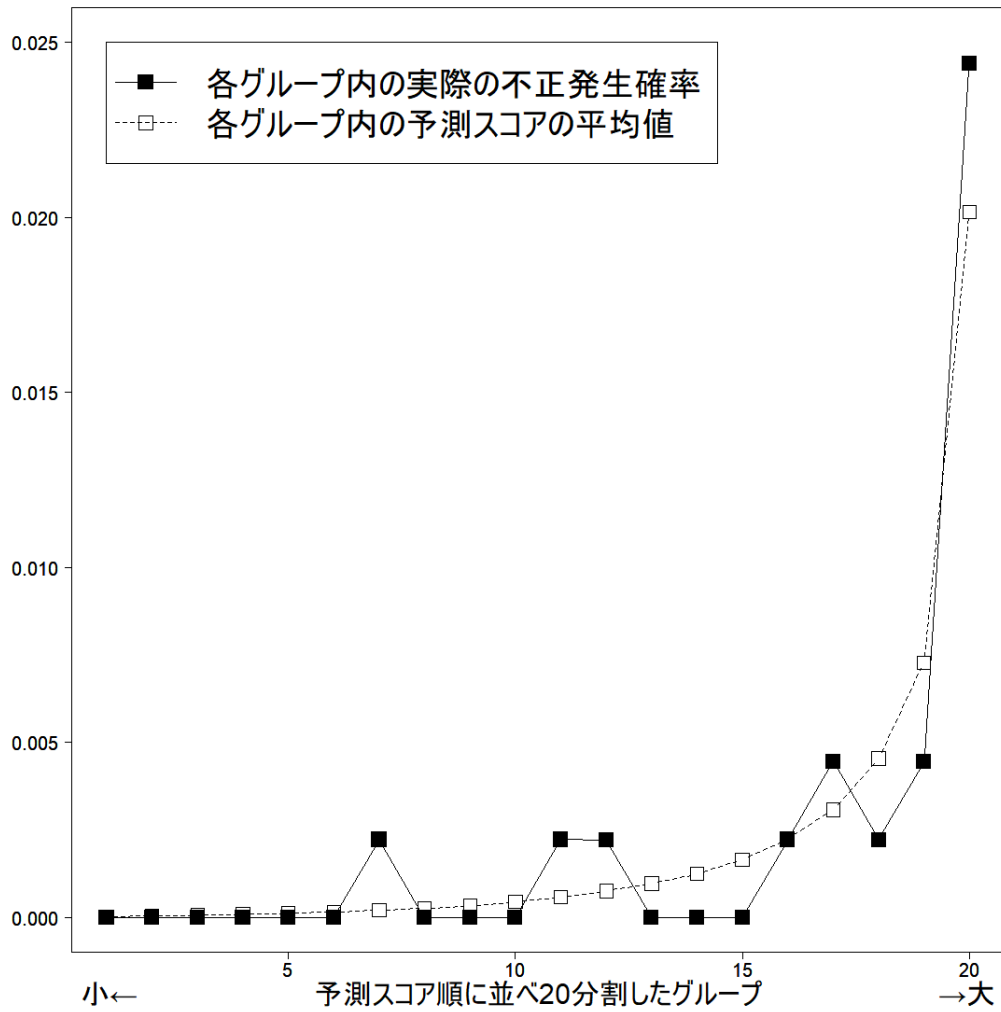


図 4

企業単位での分割（トレーニングデータ/テストデータ）によるAUC経験分布
予測 + 不正会計フラグ 2

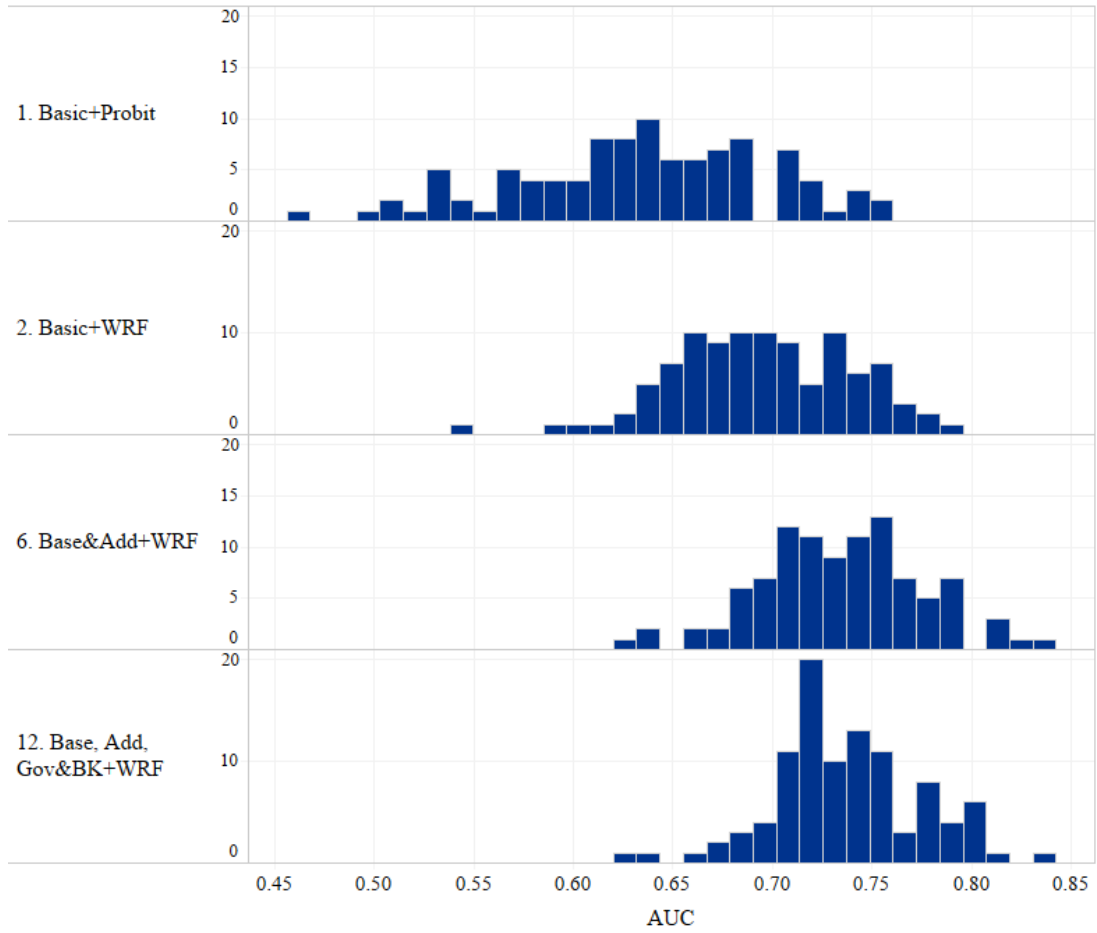
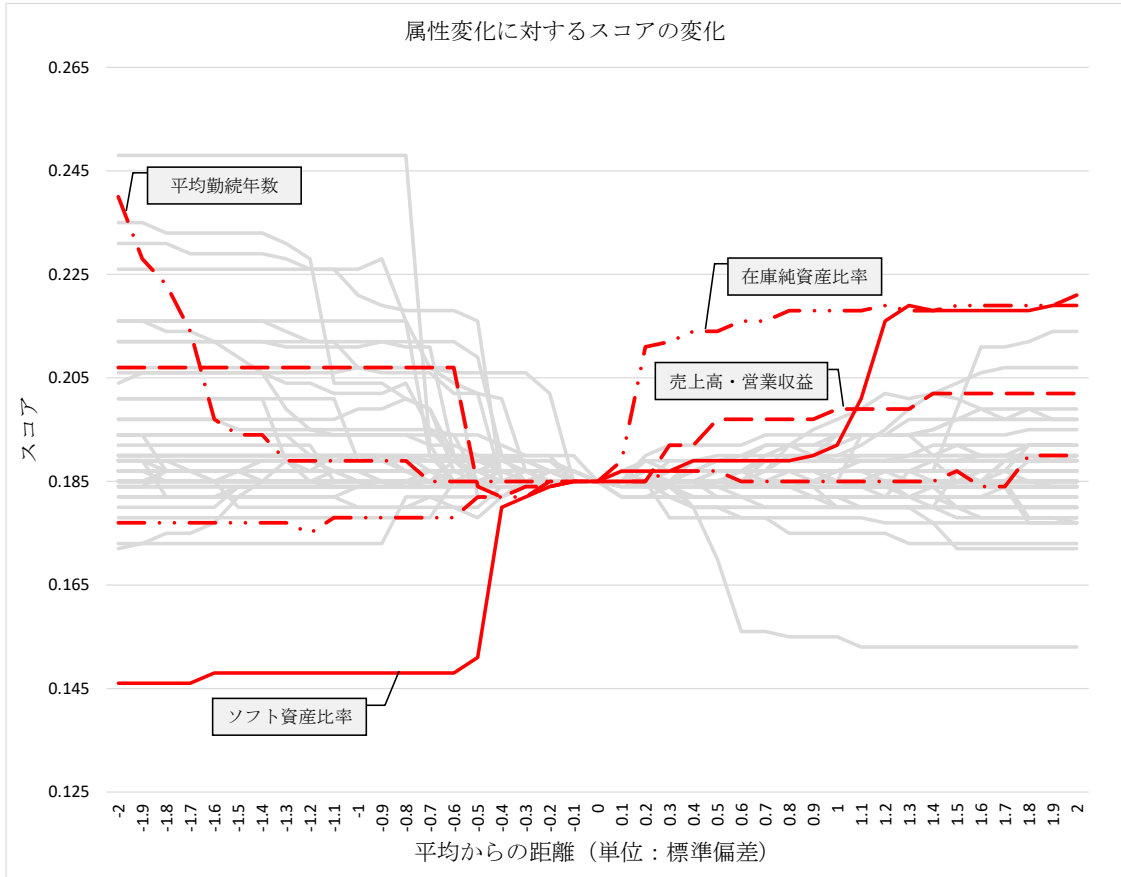


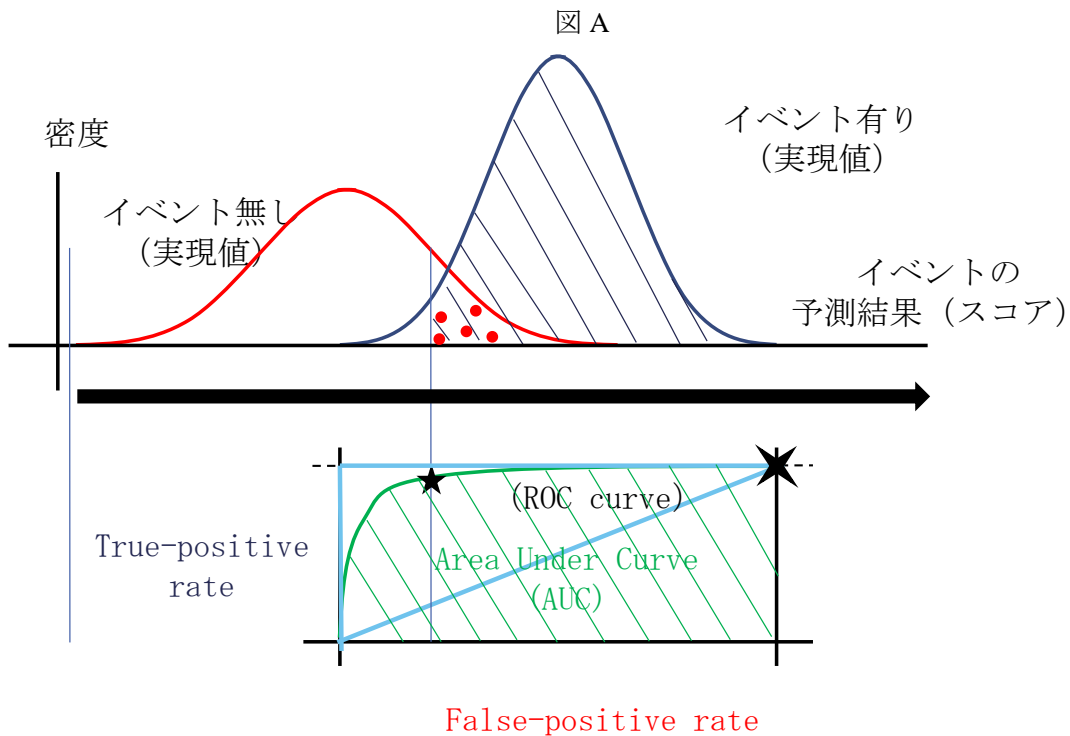
図5 平均的な属性周辺における属性変化とスコア変化の対応関係



補論1：精度指標の概念

下図 A は、特定のモデルに基づいた予測スコアと実際にイベントに直面したか否かのレコードが各サンプルに付与されている場合に、ROC (Receiver Operating Characteristic) 曲線と AUC (Area Under Curve) を用いて精度評価を行う手順を示したものである。第一に、実際にイベントが生じたか否かを基にして、予測スコアの分布を各々の群に関して描く。第二に、特定の閾値に対応する True-positive rate (実際にイベントが生じたサンプルの何割についてイベントが生じると予測されたか) と False-positive rate (実際にイベントが生じていないサンプルの何割についてイベントが生じると予測されたか) を、各々縦軸と横軸にプロットする。第三に、この点が閾値の移動に伴ってどのような軌跡 (ROC 曲線) を描くかを確認した上で、第四に、この曲線の下部に対応する面積を求めることで AUC を計算する。

この AUC はスコアに基づく分類が高い精度で機能している場合に 1 へ近づく一方で、完全にランダムな分類においては 0.5 となるため、この数値の水準を確認することで検知・予測の精度を評価することができる。実務的に期待される水準は業務の種類によって異なるものの、0.8 を超える水準が目安となるケースが多い。



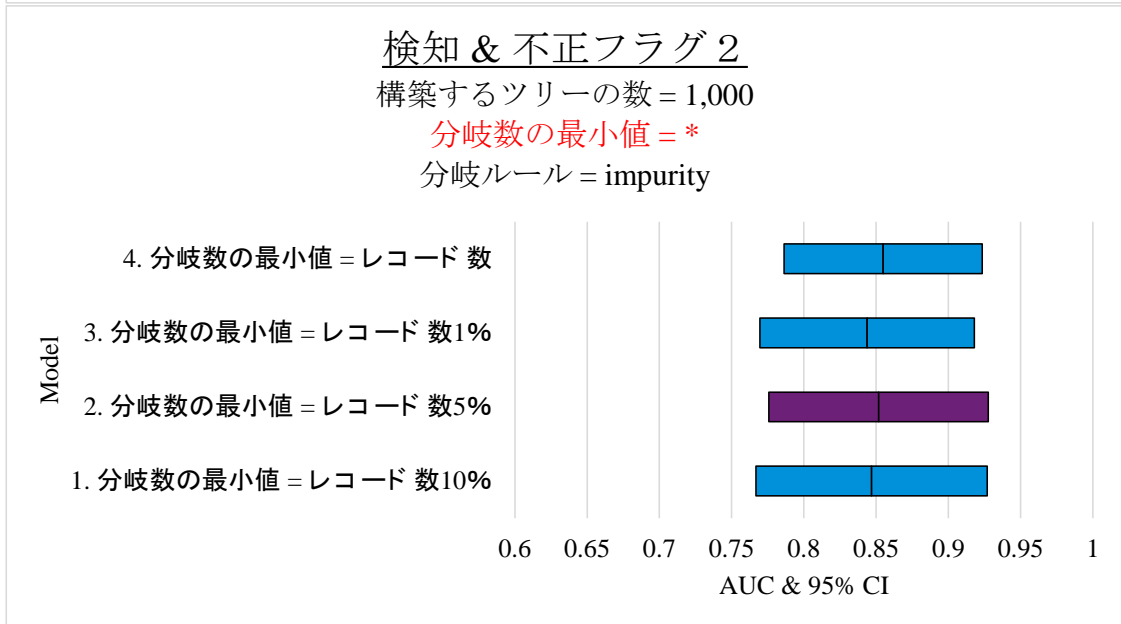
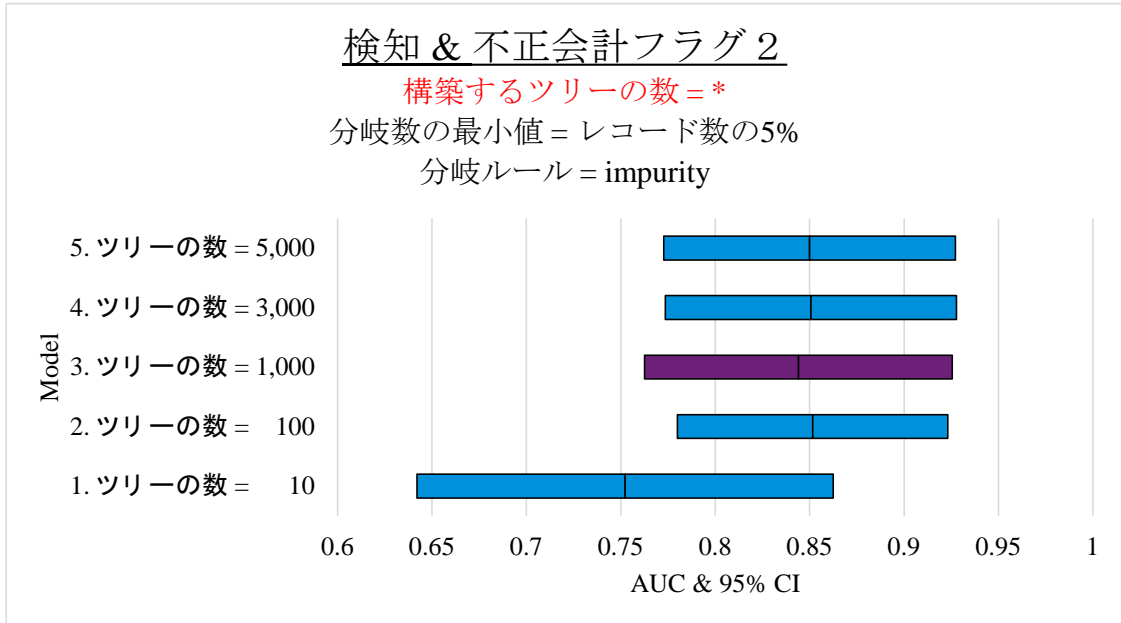
補論 2：欠損値の補完方法に関する頑健性の確認

下表は、欠損値をサンプルの平均値で補完したデータセットを用いて本稿での分析を再度実施した結果を示したものである。上段のパネルは不正会計の検知に関するモデルの精度評価結果を示しており、下段のパネルは不正会計の予測に関するモデルの精度評価結果を示している。「変更前」の列は、欠損値を0で補完した上で、別途各変数の欠損に対応するダミー変数を設定する方法をとった場合に対応している、「変更後」の列は、欠損値を平均値で補完する方法をとった場合である。平均値の算出に当たっては、企業単位で平均を算出し、企業単位でも平均が欠損取る場合には、全体の平均を算出している。

		検知 & 不正会計フラグ 1						検知 & 不正会計フラグ 2						
		変更前		変更後		差		変更前		変更後		差		
		AUC	se	AUC	se	AUC	se	AUC	se	AUC	se	AUC	se	
1	Basic	Prbt	0.71	0.06	0.71	0.06	0.00	0.00	0.73	0.04	0.75	0.04	0.02	0.00
2	Basic	WRF	0.79	0.05	0.80	0.05	0.01	0.00	0.77	0.05	0.79	0.05	0.02	0.00
3	Additional	WRF	0.84	0.04	0.83	0.05	-0.01	0.01	0.86	0.03	0.87	0.04	0.01	0.00
4	Governance	WRF	0.64	0.07	0.65	0.07	0.00	0.00	0.67	0.06	0.67	0.06	0.00	0.00
5	Bank Relation	WRF	0.68	0.06	0.62	0.07	-0.06	0.01	0.75	0.05	0.68	0.05	-0.07	0.01
6	Basic + Add	WRF	0.86	0.04	0.86	0.05	0.00	0.01	0.88	0.03	0.89	0.04	0.01	0.00
7	Basic + Gov	WRF	0.81	0.05	0.80	0.05	0.00	0.00	0.78	0.05	0.79	0.05	0.00	0.00
8	Basic + BK	WRF	0.86	0.03	0.81	0.04	-0.05	0.01	0.82	0.04	0.80	0.04	-0.02	0.00
9	Basic + Add + Gov	WRF	0.86	0.04	0.86	0.05	0.00	0.01	0.88	0.04	0.89	0.03	0.02	0.00
10	Basic + Add + BK	WRF	0.88	0.04	0.86	0.05	-0.02	0.01	0.89	0.04	0.88	0.04	-0.01	0.00
11	Basic + Gov + BK	WRF	0.86	0.03	0.83	0.05	-0.03	0.01	0.84	0.04	0.82	0.05	-0.02	0.01
12	Basic + Add + Gov + BK	WRF	0.87	0.04	0.86	0.05	-0.01	0.01	0.90	0.04	0.88	0.04	-0.01	0.01
13	Add + Gov	WRF	0.82	0.05	0.82	0.05	0.00	0.00	0.86	0.03	0.88	0.04	0.02	0.00
14	Add + BK	WRF	0.86	0.04	0.85	0.05	-0.01	0.01	0.88	0.04	0.87	0.05	-0.01	0.01
15	Add + Gov + BK	WRF	0.85	0.04	0.83	0.05	-0.01	0.01	0.88	0.04	0.87	0.05	-0.01	0.01
16	Gov + BK	WRF	0.74	0.06	0.73	0.06	0.00	0.00	0.78	0.05	0.73	0.06	-0.05	0.01

		予測 & 不正会計フラグ 1						予測 & 不正会計フラグ 2						
		変更前		変更後		差		変更前		変更後		差		
		AUC	se	AUC	se	AUC	se	AUC	se	AUC	se	AUC	se	
1	Basic	Prbt	0.68	0.07	0.69	0.07	0.01	0.00	0.65	0.06	0.70	0.06	0.05	0.00
2	Basic	WRF	0.68	0.06	0.75	0.05	0.07	-0.01	0.68	0.05	0.75	0.04	0.07	-0.01
3	Additional	WRF	0.80	0.05	0.79	0.06	0.00	0.01	0.83	0.04	0.83	0.04	0.01	0.00
4	Governance	WRF	0.64	0.07	0.64	0.07	0.00	0.00	0.70	0.05	0.69	0.05	-0.01	0.00
5	Bank Relation	WRF	0.60	0.07	0.55	0.06	-0.05	0.00	0.67	0.06	0.62	0.06	-0.04	0.00
6	Basic + Add	WRF	0.80	0.05	0.80	0.06	0.01	0.01	0.82	0.04	0.84	0.04	0.02	0.00
7	Basic + Gov	WRF	0.72	0.06	0.75	0.06	0.03	0.00	0.73	0.04	0.77	0.04	0.04	0.00
8	Basic + BK	WRF	0.73	0.05	0.73	0.04	-0.01	-0.01	0.75	0.05	0.75	0.05	-0.01	0.00
9	Basic + Add + Gov	WRF	0.80	0.06	0.80	0.06	0.01	0.01	0.82	0.04	0.84	0.04	0.02	0.00
10	Basic + Add + BK	WRF	0.83	0.04	0.83	0.06	-0.01	0.02	0.85	0.04	0.84	0.04	-0.01	0.01
11	Basic + Gov + BK	WRF	0.75	0.05	0.76	0.05	0.01	0.00	0.76	0.05	0.75	0.05	-0.01	0.00
12	Basic + Add + Gov + BK	WRF	0.82	0.05	0.82	0.06	0.00	0.01	0.85	0.04	0.85	0.04	0.00	0.01
13	Add + Gov	WRF	0.79	0.06	0.79	0.07	-0.01	0.01	0.82	0.04	0.84	0.04	0.02	0.00
14	Add + BK	WRF	0.84	0.05	0.81	0.06	-0.03	0.01	0.85	0.04	0.84	0.05	-0.02	0.01
15	Add + Gov + BK	WRF	0.82	0.05	0.81	0.06	-0.01	0.01	0.85	0.04	0.84	0.05	-0.01	0.01
16	Gov + BK	WRF	0.67	0.07	0.65	0.06	-0.03	0.00	0.72	0.05	0.69	0.05	-0.03	-0.01

補論 3 : Tuning parameter に関する結果の頑健性チェック

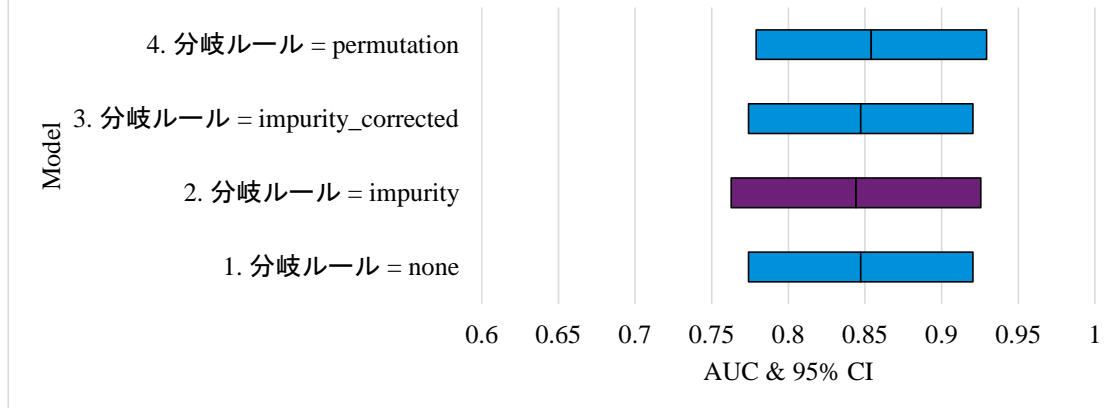


検知 & 不正フラグ 2

構築するツリーの数 = 1,000

分岐数の最小値 = レコード数の5%

分岐ルール = *



補論 4 : Tuning parameter に関する結果の頑健性チェック

上段パネルは、不正会計フラグ 2 に関する予測（縦軸）と検知（横軸）に関して各々の変数の variable importance をプロットしたものである。下段パネルは、フラグ 2 に関する予測（縦軸）とフラグ 1 に関する予測（横軸）に関して各々の変数の variable importance をプロットしたものである。

