

公開財務情報を利用した機械学習による 監査リスク分析手法に関する研究

Study on audit risk analysis method

by machine learning using public financial information

細尾 忠敬¹ 倉橋 節也¹

Tadataka Hosoo¹, Setsuya Kurahashi¹

¹ 筑波大学大学院 ビジネス科学研究科

¹ Graduate School of Business Sciences, University of Tsukuba

Abstract: The purpose of this study is to propose a new risk assessment method that complements the current audit risk assessment being conducted by experienced auditors. Specifically, we consider a method of discriminating high-risk financial figures in account units from the ones to be audited through machine learning using public financial information. From the practical view point of financial statement audits, the risk of account balances being audited is often recognized by the change in balance compared to the same period of last year and the change in correlation with other related account balances. We use, therefore, the Mahalanobis distance as the evaluation variable, a distance concept considering the correlation between variables. In conclusion, we propose a method of discriminating high-risk account combinations by clustering, an unsupervised learning method in machine learning.

1. 研究の背景

近年、上場会社の不適切な会計に関する事例は年々増加する傾向にあり、大きな社会問題となっている。例えば、ここ数年で話題になった大型の事例としては、株式会社 東芝、オリンパス株式会社等がある。一方で、金融商品取引法により求められている監査に代表される財務諸表監査の本来の目的は、財務諸表の適正性に関する結論を出すことであり、すべての不正を摘発することが目的ではない。

しかし、現実には、一般投資家には監査の過程における不正摘発の期待があり、財務諸表監査の本来の目的と一般投資家の期待との間に期待ギャップが存在しているのも事実である。このため、社会的影響が大きな不適切な会計の事例が公表された場合、翌年度以降の株主総会において、問題となった会社を担当していた監査法人から他の監査法人への変更が行われる事例が発生している。

不適切な会計によるリスクが増大する一方で、監査業務の実態については、公認会計士協会の調査によると、一社当たりの平均監査時間は増加傾向にある。また、原則として、一定の期日までに監査報告書を会社に提出しなくてはならないとする時間的な

制約が制度上存在している。さらに、東京証券取引所上場部 [1]によると、3月を決算月として採用している会社は東証全体の65.3%を占めている。その結果、決算後の監査業務は、4～5月に集中することになり、これが時間的な制約をさらに厳しいものとしている。

以上に述べた不適切な会計の事例の増加と、期待ギャップに起因する不適切な会計事例が公表された際の監査法人への影響から、財務諸表監査のリスクは増大していると考えられる。また、特に期末監査を中心に、監査の実施環境は厳しくなっており、時間的制約や人的リソースの制約が厳しい中で、監査リスクに対応することが必要になってきている。

このような監査環境に、現在の監査はリスク・アプローチに基づく監査で対応している。各監査主体は、独自に監査マニュアルを策定し、財務諸表における重要性が高い、あるいは監査リスクが高い科目や会計処理を中心に実証手続きを実施している。この監査リスクの評価はマニュアルにしたがい、一定の経験を有する監査人が判断している。

しかし、経験ある監査人が判断することが、監査上、有効かつ効率的である反面、経験があるが故の思い込みや、時間的制約等が盲点となることも考え

られる。

2. 研究の目的

本研究では、このような監査業務の遂行に対する時間的な制約や投入できる人員の制約がある中で、増大する財務諸表監査のリスクに対応するため、また、経験ある監査人の判断を補完するために、監査リスクを評価する新たな方法を検討することを目的とする。

具体的には、有価証券報告書などの公開情報を利用して勘定科目ごとに、監査リスクを分析する手法を想定している。つまり、公開情報から得られる財務数値に、機械学習の手法等を適用することにより、会社の財務数値の中からリスクが高い可能性がある数値を少ないコストでかつ機械的に判別する方法を検討する。

3. 関連研究

まず、監査業務における AI の活用をテーマとした研究としては、市原、首藤[2]がある。

市原、首藤は、分析対象とするデータにより、公開情報のみを用いて分析するマクロレベルアプローチと非公表の詳細な内部データを用いて分析するミクロレベルアプローチに分類している。さらに、マクロレベルアプローチでは、財務諸表全体としての不正会計のパターンを識別するアルゴリズム構築に AI が期待されているとし、ミクロレベルアプローチでは、異常な仕訳の検知や、勘定科目ごとの動きを他の科目との関連性から予測し異常な動きを把握する技術として AI が期待されているとしている。しかし、マクロレベルアプローチについては、先行研究についての言及、ミクロレベルアプローチについては、例示にとどめており、直接的な実証は行われていない。

次に、監査リスクをデータ分析の手法を用いて評価することをテーマとした先行研究としては、Wang, T., & Cuthbertson, R. [3]や、Cao, M., Chychyla, R., & Stewart, T. [4]がある。

Wang, T., & Cuthbertson, R.によれば、最近の分析技法の革新や視覚化、予測分析、機械学習等の技術の進歩が過去に監査データ分析の利用を妨げていた障害を取り除く可能性があるとしている。

また、Cao, M., Chychyla, R., & Stewart, T.によれば、ビッグデータ分析の手法を監査業務に活用する機会として、不正による財務諸表上の虚偽表示リスクを把握・評価する局面及びそのリスクを検証する局面等があげられている。

しかし、Wang, T., & Cuthbertson, R.の研究は、課題の提示に、Cao, M., Chychyla, R., & Stewart, T.の研究は、他分野の類似事例を紹介することにとどまっている。

さらに、不正会計の予測をテーマとした研究としては、市原・首藤[2]においてもマクロレベルのアプローチとして、先行研究がまとめられている。市原・首藤によれば、不正会計のパターンを識別するアルゴリズムに AI の様々な技術を利用することができるとし、先行研究において、不正会計予測モデルのアルゴリズムとして、多くの機械学習の手法が試されていることを論じている。

最後に、監査リスク・アプローチについては、監査基準等により定義され、その問題点の指摘については加藤[5]がある。

リスク・アプローチの考え方は、監査基準において、「監査人は、監査リスクを合理的に低い水準に抑えるために、財務諸表における重要な虚偽表示のリスクを評価し、発見リスクの水準を決定するとともに、監査上の重要性を勘案して監査計画を策定し、これに基づき監査を実施しなければならない」(監査基準、第三一項1号)とされている。なお、重要な虚偽表示のリスクは、固有リスクと統制リスクが複合的に存在することが多いとの理解から、両者を結合したリスクとして定義された概念であり、実質的には固有リスクと統制リスクを乗じたものとなる。ここで、監査基準委員会報告書 200「財務諸表監査における総括的な目的」によれば、監査リスク及び、その要素となる固有リスク、統制リスク、発見リスクは表1の通り定義されている。

表1：監査リスク及びその構成要素の定義

リスク項目	内容*	備考
監査リスク AR	監査人が、財務諸表の重要な虚偽表示を看過して誤った意見を形成する可能性	監査法人のポリシーにより所与 監査人は、このリスクを一定水準以下に抑えるように監査を計画・実施
固有リスク IR	関連する内部統制が存在していないとの仮定の上で、勘定残高等に重要となる虚偽表示が行われる可能性	経営者側に存在するリスク 勘定の特性、事業環境等の影響を受けて変化
統制リスク CR	勘定残高等に重要となる虚偽表示が、企業の内部統制によって防止又は適時に発見・是正されない可能性	経営者側に存在するリスク 期中の内部統制監査により評価
発見リスク DR	重要な虚偽表示が存在し得る場合に、監査リスクを許容可能な低い水準に抑えるために監査人が監査手続を実施してもなお発見できないリスク	監査人側に存在するリスク IR、CRの評価結果により、ARを一定水準以下に抑えるようにDRを定め、必要な手続を選択

(注) *1：監査基準委員会報告書第 200 号「財務諸表監査における総括的な目的」より抜粋し、一部修正

上記の4つのリスクの関係は、下記の式で表され、概念的には、固有リスク、統制リスク、発見リスクの3つの要素リスクを乗じたものとして監査リスクは定義されることになる。

$$\text{監査リスク(AR)} = \text{固有リスク(IR)} \times \text{統制リスク(CR)} \times \text{発見リスク(DR)}$$

また、監査業務における監査計画時には、上記の式を変形した下記の式により監査リスクを一定水準以下に抑えるように発見リスクを算定する。

$$\text{発見リスク (DR)} = \frac{\text{監査リスク (AR)}}{\text{固有リスク (IR)} \times \text{統制リスク (CR)}}$$

つまり、経営者側に存在する固有リスク、統制リスクの評価結果と、監査法人のポリシーで与えられる監査リスクを用いて、監査法人のポリシーとして定められた監査リスク水準を満たす発見リスクの水準を定め、必要な監査手続を選択する。特に計画段階では、統制リスクは前期の内部統制評価の結果に変化がないと仮定することが多いため、主に評価するのは固有リスクとなる。

さらに、加藤[5]は、監査リスクの基本評価モデルの構造から、リスク・アプローチの問題点として、以下の2点を指摘している。

まず、重要な虚偽表示のリスク(固有リスク x 統制リスク)を誤って低く見積りすぎると、本来実証手続を実施すべき場合でも、それを省略できる場合と誤って評価することで、虚偽表示を発見できない危険性が生じること、そして、会計士が重要な虚偽表示のリスクを低く見積り、実証手続が実施されないことを予見して、経営者はあえてその項目で不正を行う可能性があることである。

これらの指摘からも、監査リスクにおける固有リスク評価の重要性は高いと考えられる。

4. 研究の枠組み

研究は、勘定科目残高に対するリスクが高いと認識される状況としては、①前年同期比等による残高の大幅な変動、②他の勘定科目との相関の変化、このいずれかあるいは両方がある(森[6])との前提のもとに、ユークリッドの距離に相関とデータの位置関係による効果を足し引きした距離概念であるマハラノビス距離(救仁郷[7])を金額の変動と相関の変化をまとめて扱うことができる変数として用いた機械学習による分析が固有リスクの新たな評価手法になる可能性を検討する。

研究対象のデータとして、多くの訂正報告書が提出されている東芝の財務データを採用し、暫定的に訂正対象となった財務数値をエラーデータ、訂正後ないし訂正なしの財務数値を正常なデータとする。このデータから主要な勘定科目として、売上、売上原価、販管費、棚卸資産、営業債権、営業債務の6科目を分析対象として、各四半期の2つの勘定科目の組合せ15通り(MD01からMD15まで)について

マハラノビス距離を算定し、教師あり学習による分析として決定木による分析、教師なし学習による分析としてクラスタリングによる分析を行い、それぞれの結果について検討する。

5. 決定木による分析と考察

教師あり学習の決定木による分析では、対象とした全データを用いて、情報量基準と、ジニ基準により学習済みモデルを作成した。

まず、情報量基準の決定木により作成したモデルの正判別結果は表2の通りとなり、正判別率は82%(=47/57x100)となった。また、分岐のプロセスは、図1の樹木図の通りとなった。

表2：決定木(情報量基準)による正判別表

	エラー (観測)	正常 (観測)	周辺 度数
エラー(予測)	17	6	23
正常(予測)	4	30	34
周辺度数	21	36	57

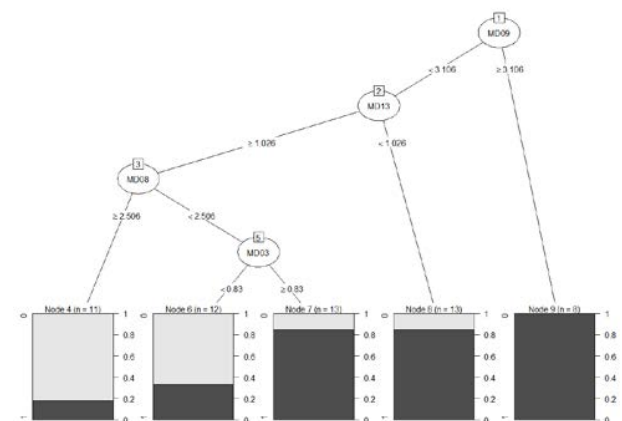


図1：決定木(情報量基準)による樹木図

また、ジニ基準の決定木により作成したモデルの正判別結果は表3の通りとなり、正判別率は74%(=42/57x100)となった。また、分岐のプロセスは、図2の樹木図の通りとなった。

表3：決定木(ジニ基準)による正判別表

	エラー (観測)	正常 (観測)	周辺 度数
エラー(予測)	12	6	18
正常(予測)	9	30	39
周辺度数	21	36	57

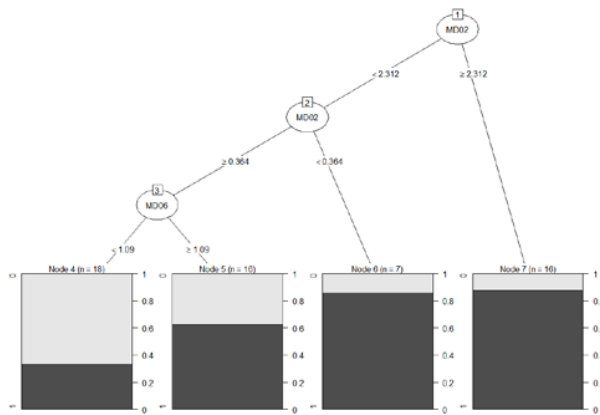


図2：決定木（ジニ基準）による樹木図

さらに、変数重要度は、表4の通りであり、情報量基準では、売上原価、棚卸資産、営業債権と関連する変数の重要度が高く、ジニ基準では、売上、売上原価と関連する変数の重要度が高くなっている。売上、売上原価、棚卸資産等の勘定科目は、東芝の不適切会計で問題となった会計処理で比較的多額の影響があった勘定科目と整合している。

表4：変数重要度の比較

マハラビス 距離	マハラビス 距離組合せ	情報量 基準	ジニ基準
MD01	売上 売上原価	14	3
MD02	売上 販管費	8	1
MD03	売上 棚卸資産	2	6
MD04	売上 営業債権	8	8
MD05	売上 営業債務	8	8
MD06	売上原価 販管費	6	2
MD07	売上原価 棚卸資産	3	4
MD08	売上原価 営業債権	4	8
MD09	売上原価 営業債務	4	4
MD10	販管費 棚卸資産	6	7
MD11	販管費 営業債権	8	7
MD12	販管費 営業債務	8	7
MD13	棚卸資産 営業債権	1	7
MD14	棚卸資産 営業債務	8	8
MD15	営業債権 営業債務	15	8

(注) 数字は重要度の順位

教師あり学習の決定木分析については、教師データを準備できる限りにおいて、比較的高い判別率のモデルが構築できるとともに、樹木図等により勘定科目単位での解釈もできることが明らかになった。

しかし、教師あり学習にはエラーデータを含む教師データが必要となるが、訂正報告書の提出状況を

考えた場合、教師データの準備に問題がある。また、教師データを反映した学習モデルを作成することから、教師データに含まれていない不正会計手法が用いられた場合の判別が難しいと考えられる。これらの点から、本研究の監査リスクの評価に利用するには問題がある。

6. クラスタリングによる分析と考察

教師なし学習のクラスタリングによる分析では、MD01 から MD15 の距離データそれぞれについて、k-means, fuzzy-c-means, 混合分布モデルの主要なクラスタリング手法を用いてクラスタリングを行い、クラスター内に含まれる距離の平均値が最大になるクラスターに含まれる変数の数が母集団の10%以下の場合に監査リスクが高い可能性があるとして仮定して解釈を行った。この仮定は、10 四半期に一度発生するかしないかの状況は、通常とは異なる状況にあり、リスクが高い可能性があるとして想定している。

まず、k-means によりクラスタリングした結果は表5の通りとなった。

表5：クラスタリング結果 (k-means)

MD	マハラビス距離 組合せ	クラスター数	CL1	CL2	CL3	CL4	CL5	計	CL1に 含まれる比率
MD01	売上高・売上原価	3	6	19	32	-	-	57	10.5%
MD02	売上高・販管費	2	18	39	-	-	-	57	31.6%
MD03	売上高・棚卸資産	3	4	25	28	-	-	57	7.0%
MD04	売上高・営業債権	2	10	47	-	-	-	57	17.5%
MD05	売上高・営業債務	2	19	38	-	-	-	57	33.3%
MD06	売上原価・販管費	2	19	38	-	-	-	57	33.3%
MD07	売上原価・棚卸資産	2	26	31	-	-	-	57	45.6%
MD08	売上原価・営業債権	5	6	4	17	18	12	57	10.5%
MD09	売上原価・営業債務	2	16	41	-	-	-	57	28.1%
MD10	販管費・棚卸資産	2	17	40	-	-	-	57	29.8%
MD11	販管費・営業債権	2	12	45	-	-	-	57	21.1%
MD12	販管費・営業債務	2	17	40	-	-	-	57	29.8%
MD13	棚卸資産・営業債権	2	7	50	-	-	-	57	12.3%
MD14	棚卸資産・営業債務	3	4	20	33	-	-	57	7.0%
MD15	営業債権・営業債務	2	22	35	-	-	-	57	38.6%

(注) 各クラスターに含まれる距離の平均値が大きいものから順に、CL1, CL2, ...と並べている。

(注) CL1~CL5の数字はクラスターに含まれる変数の個数。また、- は設定したクラスター数により該当するクラスターがないことを意味する。

同様に、fuzzy-c-means, 混合分布モデルによりクラスタリングを行い、3つの手法によるクラスタリング結果を、クラスター1 (CL1) に含まれる変数の数と母集団に対するその比率についてまとめると表6の通りとなった。

表 6 : クラスタリング結果 (総括)

MD	マハラノビス距離 組合せ	相関 係数	k-means		fuzzy-c-means		混合分布モデル	
			CL1に含 まれる値 の数	比率 %	CL1に含 まれる値 の数	比率 %	CL1に含 まれる値 の数	比率 %
MD01	売上高・売上原価	0.974	6	10.5	6	10.5	12	21.1
MD02	売上高・販管費	0.804	18	31.6	18	31.6	27	47.4
MD03	売上高・棚卸資産	-0.282	4	7.0	3	5.3	4	7
MD04	売上高・営業債権	0.679	10	17.5	10	17.5	9	15.8
MD05	売上高・営業債務	0.611	19	33.3	2	3.5	2	3.5
MD06	売上原価・販管費	0.808	19	33.3	7	12.3	12	21.1
MD07	売上原価・棚卸資産	-0.275	26	45.6	2	3.5	2	3.5
MD08	売上原価・営業債権	0.618	6	10.5	5	8.8	7	12.3
MD09	売上原価・営業債務	0.563	16	28.1	2	3.5	44	77.2
MD10	販管費・棚卸資産	-0.121	17	29.8	17	29.8	18	31.6
MD11	販管費・営業債権	0.517	12	21.1	12	21.1	14	24.6
MD12	販管費・営業債務	0.507	17	29.8	17	29.8	18	31.6
MD13	棚卸資産・営業債権	0.209	7	12.3	9	15.8	9	15.8
MD14	棚卸資産・営業債務	0.213	4	7.0	4	7.0	27	47.4
MD15	営業債権・営業債務	0.628	22	38.6	5	8.8	57	100.0

クラスタリングにより、10%以下の比率で区切った場合、固有のリスクが高い可能性があるとして判断された勘定科目の組合せは、k-means が2つの組合せ8四半期、fuzzy-c-means が7つの組合せ23四半期、混合分布モデルが3つの組合せ8四半期となった。重複を除くと、7つの組合せ24四半期となる。fuzzy-c-means 法によりリスクが高い可能性があるとして判定された組合せが最も多く、k-means 法の2つの組合せと混合分布モデルの3つの組合せがいずれも含まれている。クラスタリング手法の選択については、リスク評価の観点からリスクが高い可能性があるデータを見逃さないことを主眼において幅広く該当データを分類することができる手法が優れていると考えることができる。あるいは、さらに広くリスクが高い可能性があるデータを網羅するのであれば、3手法の結果を合成していずれかの手法によりリスクが高い可能性があるとしてきたものすべてと考えることも一つの方法である。

ここで、リスクが高い可能性があるとして判定された組合せについて、MD09 (売上原価、営業債務の組合せ) を fuzzy-c-means によりクラスタリングした結果を例として解釈を行う。図3は、MD09のクラスタリング結果の3次元散布図により示したものである。

リスクが高い可能性があるとして評価された2つのマハラノビス距離は、2007/3/31の金額の組合せと2009/3/31の金額の組合せとなっている。

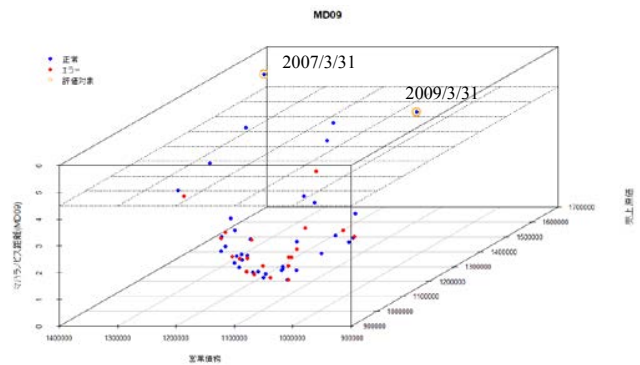


図 3 : (fuzzy-c-means) 売上原価、営業債務、マハラノビス距離の3次元散布図

詳細に分析するため、3つの2次元散布図に展開すると、売上原価、営業債務の散布図 (図4)、売上原価、マハラノビス距離の散布図 (図5)、営業債務、マハラノビス距離の散布図 (図6) となる。

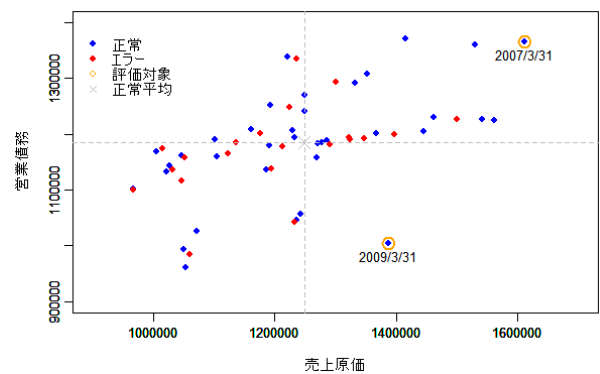


図 4 : (fuzzy-c-means) 売上原価、営業債務の散布図

売上原価と営業債務には正の相関(0.563)があるが、図4から2009/3/31の組合せは、負の相関となっており、相関に変化があることが確認できる。

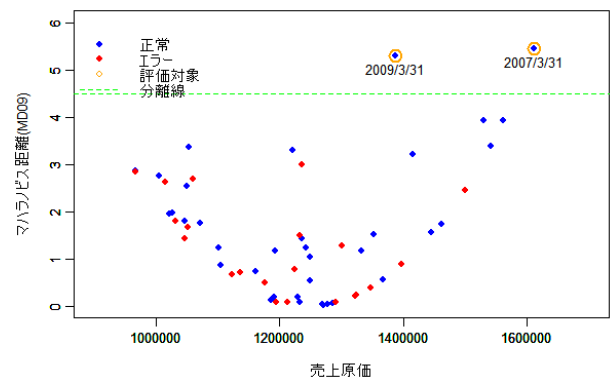


図 5 : (fuzzy-c-means) 売上原価、マハラノビス距離の散布図

また、図5から、2007/3/31の売上原価は分布の右

側に大きく外れた位置にあることがわかる。

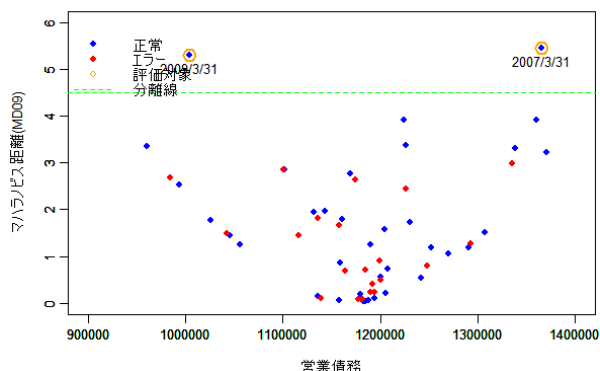


図6：(fuzzy-c-means) 営業債務、マハラノビス距離の散布図

さらに、図6から、2007/3/31の営業債務は分布の右側に大きく外れた位置にあり、2009/3/31の営業債務は分布のかなり左側にあることがわかる。

これらのことから、2007/3/31の売上原価と営業債務の組合せについては、売上原価と営業債務の両方の金額に対する固有のリスクが高く、2009/3/31の売上原価と営業債務の組合せについては、営業債務の金額及び売上原価と営業債務の相関に対する固有のリスクが高いと評価することができる。

教師なし学習のクラスタリングについては、仮説の前提となっている上記①、②の状況を平均値との比較による金額の大幅な変動、他の勘定科目の金額との相関の変化という形で明確にできる形でのクラスタリングが可能で、図4から図6により、勘定科目単位での結果の解釈性も高いことが分析の結果から明らかになった。これらのことから、本研究の目的である勘定科目単位での固有のリスクの分析に対する適合性は高いと考えられる。

一方で、クラスタリングの結果として、リスクが高い可能性があると考えられた四半期の勘定科目金額の組合せのリスク評価の妥当性を検証する合理的な基準がないことや、それに起因して、k-means, fuzzy-c-means, 混合分布モデルのどのクラスタリング手法が最適かの選択が合理的にできない等の課題がある。

7. まとめ

本研究により、機械学習の教師なし学習の手法であるクラスタリングについて、勘定科目単位の固有リスクの分析に対する適合性は高いと評価した。この方法によれば、教師データは不要であり、かつ、不正会計手法の影響を受けることなく金額の変動と相関の変化から固有のリスクが高いと想定されるデータを分類できる。さらに、監査人によるリスク評

価とやや異なる観点からリスク評価が行える点も監査人のリスク評価を補完するという意味でも有用である。

以上のことから、クラスタリング結果の妥当性の評価等に利用する合理的な基準がない等の課題はあるものの、本研究の目的である監査人によるリスク評価を補完する勘定科目単位の固有リスクの評価手法としてこの手法は有効と考えられる。

この評価手法は、財務諸表監査の実務において、監査実施初期に財務諸表監査のリスクが高い項目を洗い出す作業に有効に機能する。また、監査の最終段階で、監査完了後の財務諸表に大きなリスク項目が残っていないことを確認する作業にも有効に機能すると考えられる。

さらに、このようなリスクの分析作業についてシステム化された仕組みをあらかじめ構築することにより、監査人の人的な判断とはやや異なる観点からの監査リスクの評価を時間的制約や人的リソースの制約の中で実施することが可能となることから、財務諸表監査の質の向上に寄与することが期待される。

また、財務諸表監査以外の手法の活用として、会社の内部監査部門により行われている業務監査への適用や、金融機関における融資関連業務等への適用が考えられる。

なお、本手法の課題として、クラスタリング結果の妥当性を評価する合理的な基準がないこと等があり、今後も継続して研究を進める予定である。

参考文献

- [1] 東京証券取引所上場部. (2019). 東証上場会社 コーポレート・ガバナンス白書 2019. Retrieved from <https://www.jpx.co.jp/news/1020/190515-01.html>
- [2] 市原直通, & 首藤昭信. (2017). FinTech× 監査の現状: AI で見抜く不正会計 (特集 FinTech で変わる会計の世界). 企業会計= Accounting, 69(6), 775-783.
- [3] Wang, T., & Cuthbertson, R. (2014). Eight issues on audit data analytics we would like researched. Journal of Information Systems, 29(1), 155-162.
- [4] Cao, M., Chychyla, R., & Stewart, T. (2015). Big Data analytics in financial statement audits. Accounting Horizons, 29(2), 423-429.
- [5] 加藤達彦. (2017). 「監査の失敗」の発生原因: 実験研究でわかったリスク・アプローチの問題点 (特集「監査の失敗」の本質: ボタンのかけ違えはどこで生じているか?). 企業会計= Accounting, 69(2), 162-169.
- [6] 森實. (1994). 分析的手続監査論. 中央経済社
- [7] 救仁郷誠. (2001). マハラノビスの距離 入門. 品質工学, 9(1), 13-21.