

AIとイベントログを用いたプロセスの異常検知

アシュアランスイノベーション本部 AIラボ 公認会計士 市原直通
アシュアランスイノベーション本部 CoE推進部 公認会計士 原 誠

▶ Naoto Ichihara

2003年、当法人入社。金融機関におけるデリバティブの公正価値評価やリスク管理に関する監査、アドバイザー業務に従事。16年より会計学と機械学習を用いた不正会計予測モデルの構築・運用や監査業務におけるAI活用に関する研究開発に従事している。日本証券アナリスト協会 検定会員。



▶ Makoto Hara

主に製造業の会計監査に従事。2015年から残高確認の電子化プロジェクト、18年からプロセスマイニングの会計監査での活用プロジェクトに従事し、Digital Auditの推進に取り組んでいる。



I はじめに

EYでは監査法人と被監査会社のファイナンス部門が共創しながらデジタルトランスフォーメーション(DX)を進めることで、双方にとって新たな価値が生まれると考えています。本誌2023年新年号より連載として、監査のDXがどのように被監査会社への価値提供(リスクの適時把握やインサイト提供など)につながるかを伝えています。

本稿では、本誌23年5月号に引き続き、プロセスマイニングによる会計監査の高度化を扱います。当法人では、仕訳の異常検知を行うシステムEY Helix GLADを開発し、17年10月から運用しています。このシステムは、仕訳データから取引パターンを学習し、パターンから乖離したリスクの高い仕訳を識別します。仕訳と同様に、プロセスマイニングが分析対象とするイベントログデータについてもパターンの学習を行い、パターンから乖離したリスクの高いイベントの識別を行うことができると考えられます。これにより効率的かつ高精度な異常検知が可能となり、企業のリスク管理やコンプライアンス体制を向上させることが期待されます。AIとイベントログを用いた企業の会計プロセスにおける異常検知の方法とその利点について検討します。

II プロセス異常検知の重要性

プロセスの異常検知は不正行為や誤りを早期に発見し、リスク管理を効果的に行うために不可欠です。膨大な量のデータやトランザクションを経営者や監査人が手作業で確認することは現実的ではなく、またサンプルでの確認も有効性に限界があり、自動化された異常検知が求められています。企業は自動化された異常検知により、会計不正や人為的ミスその他、内部統制の不備などさまざまな問題を早期に特定し、対処することで、コンプライアンス体制を継続的に向上させることが可能となります。

III イベントログの役割とデータ収集

イベントログは、システムやアプリケーションが実行する操作やトランザクションに関する情報を時系列で記録したデータです。企業の会計プロセスにおいては、イベントログは販売、購買、在庫管理など、さまざまな業務活動が記録されたものとなるため、このデータを活用することで会計プロセスの検証に役立てることができます。

会計プロセスにおけるイベントログには、次のような種類があります。

- ▶ 販売・購買トランザクションログ：取引先との売買契約や受発注情報を記録するログです。
- ▶ 在庫管理ログ：商品の入出庫や在庫数の変化を記録するログです。
- ▶ 給与計算ログ：従業員の勤務、給与支払いや控除、手当などの情報を記録するログです。
- ▶ 財務報告ログ：財務諸表や税務申告書の作成・提出履歴を記録するログです。
- ▶ システム操作ログ：システムやアプリケーションの利用履歴やアクセス権限の変更などを記録するログです。

イベントログにはタイムスタンプとイベント名の他、システムによりさまざまな情報が付与されますが、どのユーザーによって入力されたのかというユーザー情報は職務分掌や適切な承認の検証の際に有用です。

イベントログの収集にはデータソースの特定、データ抽出、データのクレンジング、データの統合、データの保管、データの更新などのステップがあり、これらの手順により異常検知に適したデータを整備することができます。

IV 異常検知手法

イベントログの異常検知の手法には、異常とするデータ間の関係を事前に定めておくルールベースのアプローチと、ルールを定めずにデータの中で発生頻度が低いパターンを識別し異常とする統計・機械学習を用いたアプローチの、大きく2つがあります。前者は、検知された異常の意味が明確である一方、必要十分なルールを定めるには、監査人によるプロセスの理解が必要となります。本誌23年5月号で紹介したプロセスマイニングの活用事例は、こちらのアプローチに属するものです。後者のアプローチはアルゴリズムによってデータからパターンを学習し、そのパターンから外れるものについて検知を行うもので、イベントログのデータを利用して異常検知のモデルを構築します。イベントとイベントのつながりのパターンの学習に機械学習を用いることで、通常ではないイベントのつながりや通常ではないユーザーによる入力などを検知することができます。

V ルールベースの異常検知

異常検知に用いられるルールベースの分析として、次のようなものが挙げられます。Chiu and Jans (2019)^{※1}では、このようなルールベースの分析を欧州の銀行のデータを用いて実施した結果を示しており、同様の分析を実施する際の参考となります。

- ▶ **パフォーマンス（適合性）チェック**：プロセスマイニングを用いて、イベントログデータから実際のプロセスのフローチャートを作成し、在るべきフローと実際のフローとの差異を把握することで、想定していないフローで処理された取引を特定します。イベントの順序が想定と異なる、必要なイベントが欠落している、イベントが重複している、といった取引を検出できます。例えば、一定金額以上の発注に上位者の承認が必要とされている場合において、該当する取引全てに承認イベントが存在するかを確認することができます。承認の内容が適切であるか否かは、結果として承認が行われていることだけでは分かりませんが、承認イベントを確認することで内部統制の整備状況を評価できます。
- ▶ **職務分掌の分析**：各イベントの実施者の情報を利用して、職務分掌が適切に行われていたかを確認することができます。例えば、購買プロセスでは、一般的に注文書の承認者と物品の受領者を分けることが必要です。実際にそのような職務分掌が行われていたかを全取引について確認することができます。職務分掌の観点からは、マスターデータの変更者と、フロー上の各種イベントの実施者の関係を分析することも有用です。例えば、取引先マスターの銀行口座情報の変更と支払処理をともに行える担当者は、資金を着服できる可能性があります。さらに、このような職務分掌上の問題のある担当者が実施した他の取引を調査することも有用です。
- ▶ **取引時間の分析**：想定と異なる時間軸で処理が行われている取引を把握し、取引内容の妥当性を検討することができます。発注から検収までが長期間にわたる取引や、逆に発注後即座に検収、支払いが行われる場合等、取引時間の観点から、さまざまな異常な取引を把握することが

※1 Chiu, T.; M. Jans, *Process Mining of Event Logs: A Case Study Evaluating Internal Control Effectiveness* (Accounting Horizons 33(3), 2019) pp.141-156.

できます。また、例えば休日に処理が行われている場合等の、想定しない時間帯のイベントを特定することもできます。

いずれの分析においても、監査人がプロセスの在るべき姿（満たすべきルール）について、事前に理解した上でフローチャートを分析することが必要となります。また、実際のフローチャートを分析しながら、想定外のフローを経由した取引の妥当性を確認し、満たすべきルールを更新していく必要があります。分析対象のプロセスおよびプロセスに用いられるシステムは企業や企業内の拠点ごとに異なるため、全ての企業・拠点に一律に適用できるルールを定めることは困難です。前述の分析視点を出発点としつつ、プロセスおよびシステムに関する監査人の知識、経験を活用して、それぞれの会社・拠点に合わせた異常検知のルールを作成する必要があります。

VI 機械学習によるプロセスモデルの構築と異常検知

自然言語処理に用いられる技術の1つに、これまでの文章を踏まえると次に来る単語は何がもっともらしいかを確率的に計算し予測するというものがあります。これをイベントログに当てはめて、特定の順序で並んでいる複数のイベントを単語と見立てると、イベントログは文章のデータと捉えることができ、自然言語で用いられるモデルを適用することができます。一番単純なものはbi-gramという1つ前のイベントと後に続くイベントのペアを考慮するもので、分析に用いるデータの中で2つの連続するイベントが共起する回数に基づき、与えられたイベントの後に続くイベントの確率分布を計算します。これにより発注の起票ののちに発注の承認という2つのイベントが順番に生じるのが通常のパターンであった際に、承認ではなく納品など別のイベントが生じると通常ではない組み合わせとして検知することができます。

このアナロジーをさらに発展させてChatGPTなど

の大規模言語モデルで利用されているTransformerという技術をイベントの組み合わせに適用するという研究も行われています^{※2※3}。

また別のアプローチとして、イベントログに記録されているイベント以外の付加情報を活用するものもあります。単語のつながりである自然言語と異なり、イベントログではタイムスタンプやユーザー情報、その他の付加情報があり、これらを活用して次に来るイベントを予測することでより精度を高められる可能性があります。今回、次のようなイベントの種類がある購買プロセスを想定します。

注文書の作成
注文書の変更 (数量増加)
注文書の変更 (数量減少)
注文書の変更 (価格増加)
注文書の変更 (価格減少)
注文書の承認
注文書の却下
物品の検収
物品の検収 (マイナス)
請求書の受領
請求書の転記
請求書の転記 (マイナス)
請求書の支払い
請求書の支払い (その他)
支払いの転記
支払いの転記 (その他)

各取引のイベントごとに次にどのイベントが来るのかを予測する際に、単純に1つ前のイベントの情報のみを用いて予測を行う場合、注文書の変更や注文書の却下、物品の検収 (マイナス)、請求書の転記 (マイナス)、支払いの転記 (その他) などのイベント自体の発生が少ないものは予測ができず、そもそもまれなイベントと通常あり得ない状況の峻別^{しゅんべつ}が難しい傾向があります。

そこでユーザー ID、ベンダー ID、1つ前のイベント名、1つ前のイベントにおけるユーザー ID、ベンダー ID、1つ前からの経過時間、イベントに付随する金額などの特徴量に基づき勾配ブースティング決定木^{※4}など

※2 Bukhsh, Z.A.; Saeed, A.; Dijkman, R.M., *ProcessTransformer: Predictive Business Process Monitoring with Transformer Network* (arXiv 2018, arXiv: 2104.00721.)

※3 Moon, J.; Park, G.; Jeong, J. *POP-ON: Prediction of Process Using One-Way Language Model Based on NLP Approach* (Applied Sciences 2021, 11 (2), 864)

※4 勾配ブースティング決定木とは、分類や回帰を行う機械学習の1つである決定木という単純な予測モデルを多数組み合わせることで強力な予測モデルを構築する手法。



を用いて次のイベントを予測すると状況は改善されることが多く、注文書の変更や注文書の却下などを除き精度が大きく上がる傾向にあります。内部統制の状況にもよりますが、適切に運用されている場合には適合率（次に来るイベントがAであると予測し、実際にAであった確率）は95%を超えることもあり、一定の運用ルールに従って記録されたイベントログからある程度のルールのパターンを読み取り予測を行うことができることがうかがえます。注文書の変更や注文書の却下は予測が難しく適合率が5割～8割などばらつきますが、それでも1つ前のイベントのみしか勘案しない場合には全く予想ができなかったことを踏まえると説明力が大きく上がる傾向にあります。この予測モデルに基づき、実際のイベントが生じる確率を算出した結果が非常に低い場合、さまざまな状況を複合的に勘案して非通例な状況が生じていると考えられることから、異常として検知されることとなります。

Ⅶ 今後の検討課題

AIとイベントログを活用したプロセスの異常検知は、多くの可能性を秘めています。今後の課題も存在します。次に主要な課題と展望を示します。

1. データ

(1) データ品質の向上

高品質なデータが異常検知の精度に直結するため、データ品質の向上が重要です。システムやアプリケーションは、プロセスマイニングに適した形式でログを保有しているとは限らず、データクレンジングや前処理が必要となります。複雑な処理が必要な場合には、適切な加工が行えず、イベントデータの品質が十分でないことがあります。データの前処理に関する自動化技術の発展により、データ品質の向上が期待されます。

(2) 関連性のあるデータの収集

分析の目的に適合するイベントを特定しログデータを取得することが、分析の前提として必要です。しかし、必要なデータがシステムやアプリケーションに記録されているとは限らず、また、システム外で処理されるイベントが存在することも多くあります。データの活用を見据えて、各企業にて必要なデータを収集する体制の構築が期待されます。

(3) プロセスおよびデータの標準化および集約化

スケーラビリティの観点から、イベントログデータの標準化および集約化が重要です。日本企業では、企業ごと、さらには同一企業内でも拠点ごとに異なるシステムを利用し、データが分散管理されていることが多く、この点がボトルネックとなる可能性があります。近年のAI技術の発達によりデータ活用の効率性が高まっており、これに伴ってプロセスおよびデータの標準化および集約化への投資効果が高まっています。企業内の利用可能なデータをAIで処理できる形で集約することが、今後の競争力につながります。

2. 分析技術および分析結果への対応

(1) より高度な異常検知アルゴリズム

時系列データやテキストデータを扱う手法など新たな機械学習やディープラーニング技術の発展により、異常検知アルゴリズムの精度向上が期待されます。

(2) 人とAIの協働

異常検知の結果に対する対応策や意思決定において、人とAIが効果的に協働することが重要です。人とAIの連携がよりスムーズになるために分析結果を理解するリテラシーの向上や結果を踏まえたアクションを取る体制整備が必要となります。

Ⅷ おわりに

企業会計プロセスにおける異常検知は、リスク軽減、内部統制の強化、コンプライアンスの確保に重要な役割を果たします。会計プロセスに関連するイベントログを効率的に収集・整理し、AI技術を用いることで、効率性と精度の高い異常検知が可能となります。

AIとイベントログを活用した企業会計プロセスの異常検知は、今後の技術発展とともに、さらにリスク管理やコンプライアンスの強化に役立つことが期待されます。

お問い合わせ先

EY新日本有限責任監査法人
アシュアランスイノベーション本部
E-mail : naoto.ichihara@jp.ey.com
E-mail : makoto.hara2@jp.ey.com