

2023 年度テーマ研究論文

主査	目時 壮浩
副査	長谷川 恵一
副査	鈴木 孝則

論文題目	主題	AI による管理会計システム の展望
	副題	-予算編成 AI の開発を通じ て-

研究科	大学院会計研究科
専攻	会計専攻

学籍番号	48220014-7
氏名	内村 友紀

概要書

日時研究室が 2023 年 10 月にプライム上場企業を対象に管理会計部門の状況、課題を明らかにするために研究室で行った質問票実態調査では予算管理に AI を適用しているという回答を行なった企業は 1 社もおらず、AI を適用する予定がある企業もわずかであった。

本論文では予算編成の支援を行う LLM の構想、開発を行い、管理会計領域への AI 適用可能性や開発時の障壁となる事項について整理を行う。

AI と管理会計の可能性について考察した先行研究では、AI を適用した際に想定されるメリットやデメリットについて論じられているものの(Ulrich et al. 2022; Al-Htaybat and von Alberti-Alhtaybat 2017; Zemánková 2022), 具体的にどのような技術を用いることでメリット・デメリットが生ずるかについて検討した論文は少なかった。今日の AI は、機械学習やディープラーニング, 教師あり学習・教師なし学習・強化学習といった区分では実際にどのような処理が行えるのかが明確とはならないため, 使用するアルゴリズムやライブラリを示した上でメリットやデメリットを論じる必要がある。

そこで本論文では, 予算編成に適合する AI システムを検討するため, 予算管理の実務における問題点を明らかにし, 適合する AI がどのようなものかを探求する。結果として予算管理業務プロセスでは「予算編成にかかる期間」「予算編成にかかるコスト」「環境変化への対応」「戦略とのリンク」といった問題点を抱えており, 予算管理では「目標値の達成」「コミュニケーション」「組織内方針の統一化」に関する事項が重要視されることがわかった。

本論文で構想を検討した予算管理 LLM は Transformer 技術をもとにした LLMTime というライブラリと事前学習済み LLM によって構築される。この構造を用いることにより予算管理の実務で抱える問題点の一部を解消しつつ, 重視されている項目を阻害することなく予算編成を中心とした予算管理の支援を行うことができる。

予算管理 LLM は Base Model と予算管理 Model の 2 層構造で構築されており, Base Model は翌期以降の財務諸表数値を合理的な水準で予測するモデルであり, 企業の戦略・経済状況・市場状況といった情報を学習させることにより環境を予測した上でのモデル構築が実現する。

予算管理 Model は Base Model で予測した財務諸表数値に基づいて支援を行い, Base Model の情報に加え内部情報・他社動向・政治・顧客情報・成長性といった管理会計に必要なデータを考慮した上で予算編成数値策定の支援や予算管理行動の適正化の支援を行う。

Base Model は本研究において「有価証券報告書」「ニュースデータ」「決算短信」「企業会計基準・財務諸表規則」を用いて学習することとした。有価証券報告書にはテキストデータと数値データが混在しており, 数値データを時系列データとしテキストデータを LLM で処理するマルチモーダルなモデルにより, 財務諸表数値予測モデルの構築のための教師データとした。ニュースデータは LLMTime の仕組みによって, 有価証券報告書の数値時系列

データを補完する情報として利用する。ニュースデータの本文を利用することで、経済状況を学習し、財務諸表数値の予測や予算管理 Model での環境予測へ貢献することができる。決算短信は経営者予測情報が含まれており、財務諸表数値の予測に直接的に寄与することから教師データとし、企業会計基準・財務諸表規則は管理会計タスクに特化させるモデルを構築する上で必要な情報であると判断し、学習を行う。予算管理 LLM の構築の際には Time LLM でのファインチューニングによって事前学習済み LLM を管理会計特化型へ学習させ、RLHF を通じて Base Model や予算管理 Model を予算管理プロセスに貢献できる応答となるよう調整を行う。

予算管理 Model を構築する上で重要となる教師データは「企業の予算編成データ」「企業の予算統制データ」「予算関連議事録データ」といった企業の管理会計に直接的に関わる内部情報が重要となる。しかし、企業内部に関わるデータは入手可能性に乏しく本研究では用意することができなかった。

本研究では実際の開発プロセスを通じて、予算管理 LLM の開発上の流れの整理を行なった。また、管理会計の AI 開発の上での障壁となる事項を整理し、データ入手上の制約、データ処理上の問題、膨大な計算コストによる開発環境上の問題が起こることを明らかとし、今後の管理会計 AI 開発上の指針を示した。予算管理に AI 導入の検討を行っている企業はわずかであることから、多くの企業管理会計の導入のインフラが整っていないことが想定され、実務では管理会計 AI を導入しても運用・管理する人物がいないことが予見される。

くわえて、管理会計 AI の導入上の課題として精度検証の難しさが挙げられる。Base Model は合理的な水準で翌期以降の財務諸表数値を予測するものであるため、精度検証は可能であるが、予算編成 AI は予算管理の支援を自然言語で行うモデルであるため、長期的な業績や予算の満足度で評価するしかなく精度検証が困難である。ブラックボックスである AI の提案する管理会計手法が受け入れられるかが、今後対話型の管理会計 AI システムの普及に大きく関わってくることとなる。

目 次

第 1 章	本研究の背景と目的	1
第 2 章	管理会計の AI 活用に関する先行研究	3
2.1	管理会計の AI 適用メリット	4
2.2	管理会計の AI 適用問題点	5
2.3	管理会計の AI 適用の限界点	7
第 3 章	本研究における AI について	9
3.1	AI の概要	9
3.1.1	第 1 次 AI ブーム・第 2 次 AI ブームについて	9
3.1.2	第 3 次 AI ブームと自然言語処理	10
3.2	本研究における AI	11
3.2.1	ディープラーニング	11
3.2.2	単語ベクトル	12
3.2.3	RNN	12
3.2.4	LSTM	13
3.2.5	CNN	13
3.2.6	エンコーダー・デコーダーと Attention	14
3.2.7	Transformer	15
3.2.8	LLM	16
第 4 章	予算管理の実務に関する先行研究	18
4.1	予算管理導入企業	18
4.2	予算管理にかかる期間やコストの実態調査レビュー	19
4.3	予算管理の環境変化に関する実態調査レビュー	23
4.4	目標水準、戦略とのリンクに関する実態調査レビュー	26
4.5	予算の設定に関する実態調査レビュー	28
4.6	予算スラック・ストレッチな予算に関する実態調査レビュー	30

4.7	予算の目的に関する実態調査レビュー	32
4.8	業績評価に関する実態調査レビュー	34
4.9	予算実績差異分析に関する実態調査レビュー	37
第5章 予算編成 AI の構造と構成		40
5.1	予算編成 AI の構想	40
5.1.1	BaseModel	45
5.1.2	予算管理 Model	46
5.1.3	システムによる予算管理行動ナッジ	48
5.2	予算編成 AI のための技術	48
5.2.1	予算編成 AI の教師データ	49
5.2.2	ファインチューニング	51
5.2.3	RLHF	51
5.2.4	プロンプトエンジニアリング	53
5.2.5	時系列データと予算編成 AI	53
5.2.6	LLMTime	54
5.2.7	Meta Transformer	57
第6章 予算編成 AI の開発の流れ		59
6.1	設計の考案	59
6.2	データの収集	59
6.3	前処理	62
6.4	学習	67
6.5	チューニング等の調整	68
6.6	完成後の調整・評価	68
第7章 おわりに		70
7.1	本研究の貢献	70
7.2	本研究に残された課題	71

謝辭..... 72

参考文献..... 73

第1章 本研究の背景と目的

近年、テクノロジーの進歩により人工知能に関する関心が高まっている。筆者が早稲田図書館の提供する検索データベース WINE で 2017 年~2023 年の間でタイトルに「AI」を含む人工知能関連の論文を検索したところ 2,741,084 件ヒットした。本論文執筆現在で AI の論文の中で最も引用されている 2017 年に執筆された「Attention Is All You Need」の被引用数は 6 年間で 10 万回を超えており、AI に関する関心の高まりが見てとれる。

マネジメントや金融の領域に AI をどのように活用するかについての検討も数多く行われるようになってきているが、その一方で、管理会計分野における AI の活用については十分な検討がなされているとはいえない。昨今の AI を取り巻く環境は AI 第 3 次ブームとされ、管理会計の領域でも AI 化を検討していく必要がある。

本論文では管理会計の自動化・AI 化に関する議論の整理を行い、予算管理の実態調査から予算管理の実務上の問題点を明らかとすることで、課題解決につながる予算編成の AI の構想を行う。加えて、将来的に今後の予算管理 AI 設計時の指針となるよう予算編成 AI の構築の際に障壁となる事項の整理を行う。

テクノロジー×管理会計の領域の期待感は先行研究でも示されており、Ulrich et al.(2022)によれば、効率性と有効性の観点からデジタル化が進めば、一般的に管理会計担当者の役割が変化すると期待されており、AI を活用した予算編成について論じた Marotta and Au (2021)によれば、今後既存の予算編成アプローチに AI 手法が組み込まれ、予算編成のスピードアップ、効率性の向上により、予算編成プロセスが改善されると論じている。しかしながら、AI 化に対する批判的な意見も多く戦略的な予算の強化は意思決定者に新たな洞察をもたらす可能性を示唆しながらも、それによって生じる時間とコストの問題を指摘している。Losbichler and Lehner(2021)は、人間の合理性の欠如という観点から予測の自動化によって予測の質を高めることができるとしているが、その一方 AI システムでは経済システムのような複雑な構造は部分的にしか処理できないため AI システムでは不完全にしかモデルリングできないと主張している。Korhonen et al.(2021)によれば自動化によって代替性が高いと思われていた業務は、意外にも人間が適していると述べている。

Glazova et al.(2022)は管理会計のデジタル化によって会計とデータの調和により、人件費を増加することなく複雑なプロセスを高い専門性で実現し、ERP システム等とのシームレスな連携が期待できるとしながらも、それを実現するためには、IT インフラコンポーネントと管理会計業務プロセスのデジタル化を結合させるためのツールがないことを指摘した。

しかし、これらの先行研究は AI という多様な手法により問題解決を可能とする技術の区別を行っていない。機械学習やディープラーニング、教師あり学習や教師なし学習、強化学習レベルの粒度の区別は行われていても、具体的なアルゴリズムにまで触れて検討をしている先行研究はほとんどみられない。それゆえ、AI によって何ができるかを議論するためには、アルゴリズム等の具体的技術にまで踏み込んだ検討を行う必要がある。

そこで本論文では、執筆現在の最先端の AI を活用した予算編成 AI の検討を行う。本論文における最先端の AI とは第 3 章にて定義を行うが、主に現在主流の技術である Transformer 技術を用いた検討を行うことを指す。

また、管理会計領域のテクノロジー適用においては問題点が論じられてきているが、テクノロジー領域の進歩は目まぐるしく、執筆時点において先行研究で論じられた当時の問題点が解決されつつある問題点も存在する。特に ChatGPT をはじめとする自然言語処理モデルや画像処理モデルの技術進歩は、会計領域にも大きな影響を及ぼすことが予想される。

そこで本研究では、予算管理の有する問題点の研究を改めて整理するとともに、それぞれの問題点に対し、AI の適用可能性を検討することで、予算編成・予算管理に関わる業務の AI 化が、多くの資源や時間を割いて行われてきた予算編成・予算統制・調整などの予算機能にどのように関わるかを明らかとし、諸問題の解決に寄与する予算編成 AI の構想と繋げることとする。

第2章 管理会計の AI 活用に関する先行研究

デジタル化の急速な進展は、経営構造、意思決定、戦略の手法に大きな変化を迫っている。デジタル化は、新しいビジネスモデルを生み出し、それに伴って財務管理にも影響を及ぼしている(Zott et al. 2011)。管理会計システムは ERP システムや会計システムといった過去のデータの蓄積情報だけでなく、非会計的な管理情報システムと密接な関わりを持っている(堀井ほか 2021)。過去情報の会計情報や非財務情報がシステムを構築していることは過去データや次元間の依存関係を元に予測を行う AI に適している。さらに AI によって企業内のデータ利用の拡張も図ることができる。AI では従来のコンピューターシステムでは利用することの出来なかった非構造化データを扱うことができ、一般に世の中のデータ約 8 割は非構造化データだといわれているため、単純計算で今までの 5 倍のデータを利用した管理会計が行えることになる。しかし、Granstrand and Oskarsson(1994)が技術の多様化が必ず収益の増加をもたらすわけではないと主張していることや、自然言語処理を活用して技術の多様化が進むと企業価値にマイナスを与えることを明らかにした Matsumoto et al.(2023)などの調査から考えると、管理会計領域への AI の適用は負の影響をもたらす可能性も考えられる。

管理会計の AI 化について調査した Värzaru(2022)は管理会計における人工知能技術の受容性についてルーマニアの管理会計担当者に対して調査を実施した。結果、管理会計に AI ソリューションを適用することはプロセスの短縮化、イノベーションの発生を通じて管理会計担当者に複数の選択肢の提供を実現し、会計情報の利用度を向上させることができるとした。谷守(2019a)は予算管理と意思決定会計への AI 適用検討の結果から今後実現の期待できる管理会計ならではの機能として、「フィードフォワード・コントロール」「ビッグデータへの定性的評価分析」「最適資源配分」を挙げた。この検討結果を踏まえ、谷守(2020)は AI 意思決定のプロトタイプシステムの構想をおこなっている。このように近年では管理会計領域に AI が適用された場合の可能性について検討されており、実際の導入も構想が進んでいる。しかし、AI の導入には必ずしも良い面だけではなく、導入によって懸念される悪影響も存在する。本項にて AI 適用に考えられるメリット、デメリットそし

て AI の限界について整理を行うことで、管理会計 AI システムの構想・構築の上での可能性を検討する。

2.1 管理会計の AI 適用メリット

管理会計に AI を活用することで、どのような可能性が拡充し、期待がされているのだろうか。Warren et al.(2015)は、ビッグデータの活用によって、実務や専門職のあらゆる側面で会計を大きく変えることを指摘した。例えば財務会計の品質と会計情報の関連性について、透明性とステークホルダーの意思決定が改善され、企業報告では、基準の作成と改良が期待されるとした。PwC (2023)の調査によれば調査に参加した 85%の CEO が ERP システムと AI ツールの統合によってより効果を発揮すると考えている。Zhang(2021)は管理会計のコンピューター化は必然であり、データマイニングや機械学習の技術は品質の良い管理会計システムの構築の上では必須だと主張している。Thomas Publishing Company(2019)は機械学習が予算編成において重要な役割を担っていると主張している。例えば教師あり機械学習は、アルゴリズムを用いて企業の固定資産の寿命を監視し、修理や新規購入のタイミングを予測し、それを反映させることで直接の予算化を実現する。教師なし強化学習は、生データを評価しそこから特定の傾向を探することで、より正確な見積りでの予算編成が行えることを論じた。Marotta and Au(2021)はあらゆるソースからの情報を駆使することで同業他社の資本構造に関する情報を抽出し、比較を可能とすることやソーシャルメディアの情報を分析することで顧客ニーズの把握に繋げることができるとした。

これらの先行研究は管理会計に対して AI やデータの活用を有効に行うことで得られる可能性を推測したものであるが、実際に AI を活用した先行研究として Zemánková(2022)は Elbeltagi and Wefki(2021)が行った変数間の非線形関係を捉えることのできるというニューラルネットワークの特徴を用いてエネルギーの消費コストを正確な予測に対して、技術を応用することにより住宅不動産分野の予算編成に活用することができる可能性を示唆している。Suneja et al.(2021)はインドの MoSPI(統計・事業実施省)の大規模道路プロジェクトでは当初想定していたコストに比べ 20%超過するという問題を機械学習を用いたコスト見積りモデルによってコストの削減を成功させた。谷守(2021)は銀行業における ABC のリソースをアクティビティドライバーへ割り当てる AI 自動化システムの構築を行っており、約 98.162%の精度で自動の割り当てを可能としている。ABC を利用しない理由の

1つとして計算の複雑性や費用が挙げられており(川野 2014), このシステムは問題点克服の一要因となるだろう。Dai(2022)は会計データを分析するためにビッグデータを活用しクラウド型会計情報管理システムの構築を行った。クラウド化により, リソース共有の簡便化, データ処理の効率化を実現し, 従来財務担当者の手作業が必要だったタスクの大幅な自動化を実現した。

管理会計の AI 適用効果は谷守(2023)によれば 3 点挙げられており, ①将来志向のマネジメントによる将来思考経営, ②フィードフォワード実現によるリアルタイム経営および③経営の自動化に繋げることができるとしている。

2.2 管理会計の AI 適用問題点

昨今 AI の分野が非常に大きな注目を浴びているが, コーポレートファイナンスの領域では AI によってビジネスプロセスが大きく改善された詳細なユースケースはあまり見当たらない。予算管理の AI 利用では従来の予算管理と比較して頻繁な更新が行われることになり, 現場の状況対応力が向上されることが期待される。しかし, 作業の自動化をおこなったところで予算管理プロセスは安定せず, より高度な分析を活用しようとした際にはかかる工数, 費用が増加することが考えられる(Marotta and Au 2021)。Chen et al.(2022)は AI テクノロジーを導入した米国の企業(395 社)を対象とし AI の導入が収益増加につながるかの調査を行った。結果として自動化による業務代替を行う AI は費用の削減につながるが, 他の AI は収益性の増加効果や費用削減効果はなかったとした。Bhimania and Willcocks(2014)は会計の専門家が急速に発達するデジタル化に対応できない可能性を示唆しており, 会計活動に対して, 新たなジレンマ, 問題が発生する可能性があるとしている。会計の専門家がデジタル化に対応するためには相応の時間を必要とし, ビッグデータや新たなビジネス分析技術によって利益を得る可能性はあるが, リスクも存在していると主張している。他にもデータ, 情報, ノウハウに関する責任の所在についても課題があると論じている。

Krahel and Titera (2015)は, 会計の基準は, GAAP ベースの財務諸表が十分に動的でないため, ビッグデータの発展を反映していないことを指摘している。これは財管一致の計算書類で管理をしている企業からすればデータ活用が困難となることを示唆する。さらに財務諸表利用者が動的に抽出・検討できる生データに取って代わられる可能性を上げてい

る。

谷守(2020)では AI のブラックボックス化問題について言及している。AI は統計的アプローチとは異なり、結果を出すことを第一の目標としているため責任の所在が不明確になることで、アカウントビリティが不明確となりリスクの高い選択を AI に任せることは困難であると論じている。さらに、谷守(2021)では管理会計への AI 適用が進んでいない理由を考察しており、

- ①単なる計算処理として管理会計が考えられている可能性があり、AI よりも自動化が求められていること。
- ②定性的な判断や評価が AI では不可能だと考えられている可能性。
- ③将来予想の立て方として統計分析を利用するにとどまっていた点。

の 3 つを挙げており、事業会社が意識変革を起こさなくては管理会計への AI 適用は進まないといえるだろう。

AI 化の持つ問題はプロダクト化された AI そのものが有する問題だけでなく、AI を活用する際のデータにもあり、データに関する問題点も先行研究では指摘されている。Al-Htaybat and Alberti-Alhtaybat(2017)はデータ活用における知識の不足によって誤った意思決定につながる可能性があることを主張している。特に注意すべき状況は中途半端に知識がある状態であり、経験の浅いアナリストは異なる結果や興味深い結果がでるとそれが意味のあるものだと考えてしまう傾向にあることを懸念しており、活用の際には注意が必要である。さらにビックデータと企業報告に関する研究ではデータサイエンティスト(企業側)とステークホルダー、公認会計士の要求する情報性質の違いに関する問題も指摘している。企業報告は過去の情報として報告されることで信頼性の向上が見込めるが適時性の求められる要件では信頼性よりもリアルタイムであることが求められる。このギャップによりビックデータの会計利用時の問題が生じるのではないかと論じている。

Gärtner and Hiebl (2017)はデータが指数関数的に増え続ける中、それに対応したデータの構造化を行うことは管理会計において不可欠であると論じたが、Dai(2022)によれば現在会計関連データが構造化された形式で収集されることが基本であり、財務的な意思決定に欠かせない非構造化データの収集が十分でないことを指摘した。

Zemánková(2022)は管理会計の AI 適用においてデータの質の重要性について言及している。AI の核はデータであり、データの質は意思決定、経営に必要な質の高い情報を提供するための重要な前提条件であるとしている。近年、会計上のデータの資産化について考

察されている。

2.3 管理会計の AI 適用の限界点

メディアなどの影響で AI は万能かのように扱われることが多く、従来のコンピューターシステムの上位互換的存在と思われがちであるが、決してそのようなことはなく、利用状況に応じた使い分けが必要である。導入時、活用時の問題点とは別に AI によって実現することは不可能である限界点も存在する。例えば会計関連の業務のうち、多くは AI によって置き換えられるが Fray and Osborne (2013)の「器用さ」「創造性」「コミュニケーション」に関する業務は AI による代替は難しい(谷守 2019a)。Losbichler and Lehner(2021)は人間の合理性の欠如という観点から予測の自動化によって予測の質を高めることができるとしているが、その一方 AI システムでは経済システムのような複雑な構造は部分的にしか処理できないため AI システムでは不完全にしかモデルリングできないと主張している。谷守(2019a)は管理会計のような経済社会や将来の予測に基づいて資本や資源の最適な分配を行う業務を AI で有効に機能させることができるとしながらも、今現在の AI では創造性、芸術性、戦略性、フィロソフィーなど感情が関係する領域の業務やコミュニケーションを要する業務は困難だと論じており、その項目として、「原価企画」「経営アドバイス」「戦略策定(BSC)」「品質原価計算」「ABC /ABM」「経営計画策定・予算管理」を挙げている。

しかし、AI の代替性には限界があり業務代替的に自動化を行うことは困難であるという先行研究は「AI による自動化」と「従来のコンピューターシステムによる自動化」の性質が異なるという点を抑えた上で論ずるべきである。自動化とはプロセスやタスクを人間ではなく機械や技術にて完了させることをいい自動化によって効率性を高め、プログラミングによるルールベース化によりミスの削減が可能となる。AI 化とは自動化と比較する際においては辞書的な解釈ではなく、機能的な解釈に着目する。Wang(2019)によれば辞書的な定義とはその用語の既存の用法を要約した記述的なものであり、作業的定義とは特定の文脈においてその単語がどのように用いられるかを理解するものである。辞書的な定義でいえば「知能のいずれかであって計算機で実装されるもののうち人間が直接・間接的に設計される場合」(栗原・山川 2015)などの定義が存在するが、管理会計タスクで AI に求められる機能は「高度なデータ分析」「予測モデリング」「データ統合支援」「意思決定のサポー

ト」「異常検知」「プロセスの最適化」などが考えられる。これらの機能は業務代替的で自動化に密接に関わるものもあるが、予測タスクといった自動化に関連しないものも存在する。また、同じ自動化であっても従来のコンピューターシステムとは性質が異なる。AIによる自動化とは AI が学習を通じてタスクの判断を行う自動化でありルールベース化された自動化に比べ高度なタスクに対応できるが、ミスが生じる可能性もある。

例えば価格設定業務は機械よりも人間の方が適している(Korhonen et al. 2021)との主張があったが、従来の自動化では自動化すべきでないとされていた事項であっても AI による自動化であれば人間よりも優れた成果を出すことも考えられる。

AI であれば多量の購買データや顧客行動を考慮した上での価格設定がなされるため、ルールベースでは対応できなかった自動化が実現される。以前までの AI はルールベースでしか対応できず、価格設定などの定性的業務への対応は難しかった。そこで次章にて AI の変遷を辿るとともに本論文における AI の定義づけを行う。

第3章 本研究における AI について

3.1 AI の概要

管理会計の領域において AI 分野は馴染みがないものであり、管理会計に AI を適用させるための主張を解釈するためには AI の知識がある程度なくてはならない。そこで本項では予算編成 AI のために用いた AI を理解するための知識の整理を行う。

一意に AI といっても AI という言葉に含まれている意味は非常に多い。AI という言葉の意味は多くの研究者が定義していることがその表れである。AI という言葉を明確に定義することは困難であるが、本論文における AI の考えを明確化することで議論の際に誤った方向性の解釈を避けることができる。そこで本章では AI の歴史の変遷から始まり、本研究で主に検討を行う AI を明らかにすることで、管理会計における問題点をどのような技術で解決に繋げるかを明確化することができる。

3.1.1 第 1 次 AI ブーム・第 2 次 AI ブームについて

AI の歴史は 1950 年代から続いており、その間技術進歩により AI によって可能なことが大きく変わっていることが要因の一つである。人工知能という言葉は 1955 年に John McCarthy によって定義された。1957 年に Frank Rosenblatt がパーセプトロンというモデルを開発すると AI に注目が集まった。パーセプトロンとは入力(特徴)を受け取り、それらを重み付けし、合計して、特定の閾値を超えた場合に出力(決定)を行う単純なモデルであり、線形分離可能な問題に対してのみ有効であった。この技術は現代のニューラルネットワークやディープラーニング技術への重要な 1 歩であり脳のニューロンを模倣することの取り組みを開始した。1967 年には ELIZA という自然言語処理プログラムが開発された人間とコンピュータ間の対話を模倣することを目的としておりのちの会話 bot や自然言語 AI に影響を与えることとなった。1980 年代半ばまでは AI ウィンターと呼ばれる時期に突入り第 1 次 AI ブームは終焉したが、第 2 次ブームとしてエキスパートシステムと呼ばれる AI が再度熱をつけた。エキスパートシステムとは特定の領域の専門性を持つ知識ベースを

有したシステムでありこの知識はルールベースで定義される。知識ベースに格納された情報を利用して新しい事例に対する回答を算出するシステムであった。しかしながら、エキスパートシステムは汎用性にかけており期待された通りの性能が出なかったことや開発・維持コストが高かったことから関心が失われ第2次 AI ブームは終焉した。2000 年代初頭に入ると第3次 AI ブームに突入し大きく火をつけた要因となったのはディープラーニングの進化である。

3.1.2 第3次 AI ブームと自然言語処理

Krizhevsky et al.(2012)は ImageNet LSVRC-2012 コンテストでディープラーニングに基づく CNN (Convolutional Neural Network) (Cun et al. 1990)を用いた手法 AlexNet で優秀な成績を納め衝撃を与えた。このコンテストは 120 万の高解像度画像を 1000 の異なるクラスに分類しエラー率を競うコンテストであり、CNN を用いた技術によってエラー率を減少させ訓練速度を大きく改善した。AlexNet の登場によりディープラーニングに大きな関心が集まり発展に大きな貢献をもたらした(Czum 2020)。ResNet という手法が登場した LSVRC-2015 コンテストでは画像識別精度が人間の精度を超えることとなった(He et al.2015)。

AI は約 70 年の歴史を経て多様な技術が開発されたが、その結果 AI という用語に包括される技術が非常に多岐に渡るものとなった。現代における AI という言葉の文脈でパーセプトロンを意図しているということは稀であるが、機械学習アプローチとディープラーニングアプローチは共に AI という言葉が使われる。機械学習とディープラーニングはともにデータから学習して予測や分類を行う技術であり、ディープラーニングは機械学習の一部であるが、アプローチや出力結果には大きな違いがある。機械学習の伝統的なモデル(決定木、線形回帰、ロジスティック回帰など)は、統計的な推論に根差した出力を提供する。結果として特定の予測がなされた理由を理解しやすく、BI や医療分野など、意思決定の根拠が重要とされる領域で重宝される。モデルの複雑性と解釈可能性はトレードオフとなるが SHAP や LIME といった説明可能 AI も活用可能なレベルにあり、ビジネス分野でも複雑なモデルよりもシンプルで AI の意図が把握しやすいモデルが用いられるケースが多い。比べてディープラーニングはその構造の複雑さと多層性により、高い予測性能を達成しやすいが、その結果の解釈は困難である。ディープラーニングは「ブラックボックス」

と見なされることが多く、ネットワークがどのように特定の予測に至ったかを理解することは非常に難しい。このアプローチの主な目的は、予測精度、パターン認識の能力、そして新しいデータに対する適応性を最大化することであり、特に画像や音声の認識、自然言語処理などの領域で顕著な成果を上げている。他にもディープラーニングには機械学習と比べ大量のデータを必要とする点などから活用にコストや手間がかかることが多い点の特徴として挙げられる。

直近の AI 分野の技術でブレイクスルーをおこすのはディープラーニング絡みの技術が多く、近年話題となっている生成型 AI は多くがディープラーニングを活用した Transformer 技術(Vaswani et al. 2017)に関連していることが多い。

3.2 本研究における AI

本論文では最先端の AI 技術を活用した管理会計を検討する。具体的には自然言語処理技術からなる Large Language Models(以後、LLM と記載)を用いた予算管理支援システムの構築を検討し、管理会計の AI 適用可能性について考察する。そこで本論文における最先端 AI が何を指すかを明確化するため、近年の自然言語処理に関する技術の整理を行う。

3.2.1 ディープラーニング

ディープラーニング(Rumelhart et al.1986)とは、計算モデルにおける近似の概念を中心に据えた、機械学習手法である。本手法は、特に大規模なデータセット上で複雑なパターンや関係性を抽出し、それらを数学的に扱いやすい形で近似する能力において顕著な性能を示す。多層構造を持つニューラルネットワークを用いて、入力データから段階的により抽象的な特徴を学習し、これによりデータの本質的な特性を捉える。このアプローチの根底にあるのは、データの内在する複雑性を、より単純化された数学的表現に変換することである。例えば画像のタスクにおいて浅い層では基本的なデータ特徴(例：エッジ、色相)が捉えられ、深い層に進むにつれて、これら基本特徴の組み合わせから形成されるより複雑な抽象的特徴(例：物体の形状、相互関係)が学習される。このプロセスは、データの多様性と複雑性を統合し、実用的な予測や分類に必要な本質的な情報を抽出する。

ディープラーニングの成功は、主にその優れた近似能力に依存する。この近似は、ネットワークが訓練データから得た知識を基に、未知のデータに対しても適用可能な一般化されたモデルを構築することに他ならない。しかし、この近似プロセスの効果を最大化する

ためには、大量の訓練データが不可欠である。データ量が増えるほど、ネットワークはより精確な近似を生成し、結果としてより高い精度の予測や分類が可能となる。

3.2.2 単語ベクトル

自然言語処理とは機械にテキスト情報を読み込ませて処理することであるが、機械は文字列のままでは意味を認識することができない。そこで機械にテキスト情報を理解させるための工夫としてテキストの単語ベクトル化(トークン化)処理を行う必要がある。

単語ベクトルとは、自然言語処理における根幹的な概念であり、単語の意味的属性を固定長の実数ベクトルにエンコードする手法である。この手法は、単語間の意味的な類似性や文脈依存性を数値的に捉え、複雑な言語データを機械学習モデルが扱いやすい形に変換する。単語を多次元空間上の点としてベクトル表現する。ベクトルは、単語の意味的特徴を捉え、単語間の関係を定量化する。大規模なテキストコーパスを単語の文脈内での使用法に基づいてその意味を捉える学習方法が近年の主流である。以前までの単語表現は、機械学習や統計モデリングにおいて広く用いられていた。これにはワンホットエンコーディングや TF-IDF などの伝統的な手法が含まれ、特にデータが限られた場合や特定の統計モデルに適用する際に有用である。ディープラーニングに基づく自然言語処理の分野で一般的に使用される単語表現を単語分散表現といい、Word2Vec, GloVe, FastText のような手法は、単語を低次元で密なベクトルとして表現し、これによって単語の意味的な類似性や文脈依存性を捉える。これらの手法は、大規模なテキストデータセットを使用して訓練され、単語間の複雑な関係を学習する。単語ベクトルの導入は、特に RNN(Recurrent Neural Networks) (Rumelhart et al. 1985)などのシーケンスモデリングにおいて重要な意味を持つ。

3.2.3 RNN

RNN は、時間的な順序を持つデータを処理するために設計されたネットワークであり、単語の意味的な連続性を捉えるのに適している。単語ベクトルを用いることで、RNN は各単語の豊かな意味的特徴を入力として受け取り、より複雑な文脈や言語構造を学習することが可能となる。RNN の核心は、過去の情報を記憶し、それを現在の入力と組み合わせて処理する能力にある。この「記憶」は、ネットワークの隠れ層における内部状態によって

実現されている。各時点での入力、前の時点の内部状態と組み合わせられ、出力を生成する際に過去の情報を考慮することが可能となる。RNN は、データの時間的な側面を捉えることにより、データの内在する複雑性をより豊かな文脈で近似し、解析することが可能になる。例えば、言語における単語の意味は、その単語が現れる文脈に大きく依存するが、RNN はこのような文脈を捉えることで、より正確な言語理解を可能にする。この能力は、ディープラーニングが複雑なパターンを近似する上での多様性を示す。単語ベクトルと RNN の組み合わせは、単語の意味的特徴を深く捉え、文脈内での解釈を可能にする。特に、文の生成、機械翻訳、感情分析、質問応答システムなど多様な NLP タスクにおいて、この組み合わせは顕著な成果を示す。RNN は、単なる単語の並びを超えた文脈的な理解を達成し、自然言語の複雑さを機械的に処理する能力を大きく向上させている。

しかし、RNN は長いシーケンスを扱うことを苦手としておりその問題を解決するために生み出された技術が LSTM(Long Short-Term Memory)(Hochreiter and Schmidhuber 1997)である。

3.2.4 LSTM

LSTM は、特に RNN が持つ短期的な依存関係の制約を克服し、より長期的な依存関係を捉える能力を有する。LSTM の核心的な特徴は、忘却ゲートを通じて情報を選択的に保持または忘却する能力にある。このゲートは、過去の隠れ状態と現在の入力に基づいて、隠れ状態に格納される情報を制御する。入力ゲートは新しい情報をどの程度隠れ状態に加えるかを決定し、セル状態は長期的な依存関係を保持するメカニズムを提供する。出力ゲートは、セル状態の情報をどの程度次の層に渡すかを決定し、現在の出力を生成する。

3.2.5 CNN

CNN は、主に画像処理やコンピュータビジョンの分野で広く用いられるが、自然言語処理においてもその応用範囲を拡大している。CNN は、ローカルな特徴を効率的に抽出し、それらを高次の表現に統合する能力を持つ。NLP における CNN の応用は、テキストデータに対しても局所的なパターン(n-gram など)を抽出することにある。このアプローチは、特に文書分類や感情分析などのタスクで有効である。CNN は単語やフレーズレベルでの特徴を捉え、これらを組み合わせて全体の文脈を理解する。LSTM がシーケンスデータの

長期的な依存関係をモデル化するのに対し、CNN は局所的な特徴の抽出に優れる。自然言語処理における CNN の利用は、特にテキストの局所的なパターンを捉えることに重点を置いており、異なる種類のタスクに対して多様なアプローチを提供する。このように、CNN は自然言語処理の分野においても重要な役割を果たし、言語データの分析と理解の幅を広げている。

3.2.6 エンコーダー・デコーダーと Attention

RNN や CNN 技術は 2000 年初頭頃から自然言語処理の主流であり発展を支えてきた技術であるが、現在ではリソースが限られている状況や小規模なデータセットを用いる場合を除いて利用されることは少なく主流とはいえない。2017 年に公開された Transformer 技術によって多くの役目が奪われた。ここでは現在の自然言語処理の主流である Transformer を説明するにあたって、「エンコーダー・デコーダー」「Attention」といった仕組みの紹介を行う。

エンコーダー・デコーダー構造における Attention メカニズムは RNN や畳み込みニューラルネットワーク CNN と比較して、自然言語処理の分野で大きな進歩をもたらした。エンコーダー・デコーダー構造は、特に機械翻訳や自然言語理解の分野で重要な役割を果たす。この構造は、2つの主要なコンポーネント、エンコーダーとデコーダーから成り立っており、一般にシーケンスからシーケンスへの変換タスクに用いられる。

エンコーダー・デコーダー構造の特徴エンコーダーでは、入力シーケンスを受け取り、それを固定長のコンテキストベクトルに変換する。このコンテキストベクトルは、入力シーケンスの全ての重要な情報をエンコードし、デコーダーへ渡される。デコーダーはエンコーダーからのコンテキストベクトルを受け取り、目的のシーケンス(例:機械翻訳における翻訳後のテキスト)を生成する。エンコーダー・デコーダー構造は、入力と出力が異なる形式や意味を持つ複雑なタスクに適しており、自然言語処理の分野において広範な応用が可能である。エンコーダーが入力データの意味を捉え、デコーダーがこの情報を基に新たな出力を生成するこの一連のプロセスは、言語変換や情報抽出の問題を解決する上で非常に効果的である。

Attention は、入力データの重要な部分に「注意(注目)」し、それに応じて出力を生成する機能である。Attention メカニズムは、入力シーケンスの中で特に重要な部分に焦点を当

て、その情報を強調する。これにより、デコーダーは関連性の高い情報に基づいてより正確な出力を生成が可能となる。Attention は RNN や CNN に比べ長いシーケンスへの対処に適している。RNN はシーケンスデータを逐次的に処理し、過去の隠れ状態に基づいて新たな状態を生成するが、Attention はシーケンス全体を直接参照し、任意の 2 点間の関連性を直接モデル化する。その結果 RNN では長いシーケンスによって情報の喪失や勾配の消失問題が発生しやすい。一方、Attention は各時点で全シーケンスにアクセスし、重要な情報を直接活用するため、シーケンスの全体的な文脈を考慮し、局所的な特徴と全体的な構造の両方を統合することができる。よって、Attention メカニズムは、自然言語処理の分野における RNN や CNN の制約を克服し、特に長距離の依存関係や複雑なシーケンス構造のモデリングにおいて大きな進歩をもたらしている。このメカニズムによって、NLP タスクにおけるモデルの性能が大幅に向上し、より洗練された言語理解が可能となったため自然言語処理のブレイクスルーを果たした技術である。

3.2.7 Transformer

近年の AI、特にディープラーニングの領域は Transformer 技術によって大きな発展を遂げた。Transformer は元々翻訳のために生まれた技術であったが汎用性に優れており、2023 年現在では様々な AI モデルが Transformer ベースとなっている。Transformer 登場以前の自然言語処理は RNN や CNN を用いたエンコーダーデコーダーモデルであり、Attention 機構によって結合されているモデルが主流であった。Transformer では RNN や CNN を用いずに Attention 機構を組み合わせ Multi-Head Attention とし性能の向上を果たした。Attention とは出力シーケンスと入力シーケンスにおける長距離依存関係をモデリングすることができるモデルである。Transformer モデルは、埋め込みされた単語ベクトルを、Attention を並列処理で行う Multi-Head Attention で処理しシーケンスのモデリングを行う。それに対し、Add & Norm により残差接続や正則化を行い Feed Forward 層により Multi-Head Attention にて蓄積した知識を非線形へと変換し、複雑なパターンを学習する能力を得る。この階層を複数回繰り返し出力層では Softmax で次単語の生成確率を計算する。

本論文における最先端の AI とは現在多くの AI タスクにおいて主流となる Transformer

¹ RNN, CNN は現在でも一線級で用いられており、Transformer 技術より高い精度を実現する領域も存在する。(Zhao et al. 2021)

モデル以後に提案された AI をする。Transformer モデルであれば以前は不可能とされていた管理会計 AI タスクも解決する兆しが見え、今後の管理会計 AI の発展を支える技術となりうると考えられるためである。経済状況、市場状況、自社状況といった複雑な要素が混在する管理会計で Transformer モデルを用いることは、その状況を適切にモデリングできる可能性を有しているといえる。

しかしながら、Transformer ゆえの問題点も存在しており、これらの問題点が生じる可能性を留意する必要がある。

①固定長の入力

Transformer モデルは固定長の入力しか扱えないため長文に対する適用能力が不足している。この特性により古い情報が切り捨てられ、管理会計業務プロセスの論理立てに対応できない恐れがある。

②Attention の計算コスト

Self-Attention では全てのトークンの位置関係を計算するためやりとりが長期化すると計算コストの観点から全てをモデリングすることが難しくなる。

③位置エンコーディングの限界

位置エンコーディングも固定長での設計となっているため、長文の後半部分ではエンコーディング数が不足し、情報が適切に読み取れない可能性がある。

これらの問題点はいくつかの解決案が研究されてきてはいるが(Guo et al.2022;Zaheer et al.2020)依然として管理会計 AI では課題となることが予見される。Transformer 技術を用いて世の中に大きな影響をもたらした製品として Chat GPT(Bahrini et al. 2023)があるが、Chat GPT は LLM というモデルによって構築されている。本研究では汎用的タスクへの対応性の高い LLM を用いることで複雑な要素の噛み合う管理会計の諸問題への支援を図ることのできる AI を構築することが可能となる。

3.2.8 LLM

本論文で検討する AI システムは LLM にて予算管理業務活動の最適化を行うための支援システムである。ここでは LLM のモデルについての説明を行う。Language Models とは単語の系列がどの程度発生しやすいかをモデル化したものであり、単語の系列 $(x_1, x_2 \dots x_L)$ にその生成確率である $p(x_1, x_2 \dots x_L)$ を割り当てる確率モデル p である。

² $p(x_1, x_2 \dots x_L)$ を連鎖律で分解すると $p(x_L|x_1, x_2 \dots x_{L-1})$ となり、自己回帰言語モデルと呼ぶ。これを言語で当てはめる場合、以下のように「吾輩は」の後に続く言葉を argmax で確率モデルとして表している。

$$p(\text{猫}|\text{我輩, は})=0.8$$

$$p(\text{犬}|\text{我輩, は})=0.02$$

$$p(\text{売上}|\text{我輩, は})=0.001$$

したがって、Language Models とは過去の単語列から次の単語を予測するモデルであり、Transformer ではこれをエンコーダー、デコーダー、Attention 機構によって高精度なモデルを実現している。関数近似値である次の単語を予測する Language Models であるが、多くの言語パターンを学習させることによって、精度の高いやりとりが実現する。LLM とは大規模なデータセットを学習させた Language Models であり、本論文では既存のオープンソース LLM に管理会計情報を学習させることで、管理会計タスクを支援する特化型 LLM を構築することが可能である。特化型 LLM を構築する上では実務における問題点の解消となるモデルを検討することが必要であり、実務上の予算管理における課題を把握することでモデルに学習させるデータの見当がつくようになる。

² 連鎖律:ある事象 A とある事象 B が同時に起こる確率は事象 A が起きた条件下での事象 B が起こる条件付き確率の積として表すことができる。よって、 $p(A,B)=p(A) \cdot p(B|A)$ に分解可能である。

例: $p(x_1, x_2, x_3)$ を求める場合

① $p(x_1)$ の計算 ②. $p(x_2|x_1)$ の計算 (x_1 の条件下で x_2 の条件付き確率)

③. $p(x_3|x_1, x_2)$ の計算(x_1 と x_2 の条件下での x_3 の条件付き確率) ④ これらを掛け合わせることで全系列の同時確率

第4章 予算管理の実務に関する先行研究

本章では予算管理の実務に関する先行研究を整理する。予算管理の AI 化、自動化といった技術の適用を行う上で、実務上どのような予算管理の利用が行われているのかを明らかにする必要がある。予算管理の問題点、課題点を整理するとともにどのような予算管理が企業にとって正の影響をもたらすのかについての考察を行うことにより、目指すべき予算管理システム像を明確にする。よって、先行研究での実態調査から企業が望むシステムのニーズの抽出を行い、現在企業が抱えている課題や利用している手法が自動化、AI 化の観点からアプローチ可能かについて検討していく。

先行研究において管理会計や予算管理の実態調査が実施されているため、調査結果を比較し、予算管理に対する管理会計担当者の意識の推移を推測することで、予算管理のシステムで求められている要素を明らかにする。予算管理領域は原価計算領域の実態調査に比べて業種間の差が少ない(榎谷ほか 2022)ため、多様の研究者の調査結果を同列に捉えて比較を行っても問題ないと考えられる。

実態調査の先行研究では AI という視点から実態調査を行なった研究は存在しないものの、実態調査をもとに AI を導入した際のアプローチを検討し、あるべき予算編成 AI の構想を行うことで実務的問題点を補うことの出来る予算編成 AI システム像を明らかとすることができる。³

4.1 予算管理導入企業

予算管理は非常に多くの企業で導入されており、清水ほか(2021)の調査では、一部上場企業で予算管理システムを実施している企業は 178 社中 168 社であり、予算管理システムを有していない 10 社においても、うち 9 社は財務的な計画のためのシステムを有していることがわかっている。横田・妹尾(2011)の調査でも予算管理を実施している企業は 98.6%

³ 今回参考とした先行研究では 7 点尺度のリッカートが主流であったため、特段記載のない場合は 7 点尺度のリッカートとする(5 点尺度の場合別途記載)

であり、10年以上前の時点から企業において予算管理は重要だと考えられており、今現在においても重要性は低下していないと認識することができる。

図表 1 企業で予算管理の実施

	予算管理を実施している	予算管理を実施していない	その他
清水ほか (2021)	168 社(94.4%)	1 社(0.6%)	9 社(5.1%)
清水(2016)	188 社(88.9%)	2 社(1%)	10 社(10.1%)
横田・妹尾 (2011)	218 社 (98.6%)	2 社 (0.9%)	1 社 (0.5%)

出典：筆者作成

予算管理が多くの企業で実施されていることは AI 構築の観点からも望ましいことである。(Gärtner and Hiebl 2017)が指摘したように管理会計の AI 活用においてデータの収集は非常に重要である。予算管理を利用している企業が多いということから企業にデータ収集の機会があり、多くの企業の管理会計データが必要となる管理会計 AI の活用に重要な要素となる。

4.2 予算管理にかかる期間やコストの実態調査レビュー

予算管理の問題点についての問いで特に課題とされた項目の1つは予算編成の時間的負担やコストに関する項目である。清水ほか(2021)の調査で「予算は手続が複雑で多額のコストがかかっている」という問いでは平均値 4.72, 7 点 7.9%, 6 点 25.8%, 5 点 27.5%となった。1 点と回答した企業は 0.6%しか存在していないことからほぼ全ての企業が少なからず予算は手続が複雑で多額のコストがかかっていることがわかる。他にも、吉田ほか(2019)の調査では「予算管理の時間的負担」に対する問いで製造業平均値 5.08, 非製造業平均値 4.36 と 2021 年の調査と遜色ない結果となっている。

具体的にどの程度企業が予算編成に時間を費やしているかの調査では上東(2021)によると大規模企業では 4 ヶ月 12.9%, 3 ヶ月 39.7%, 2 ヶ月 29.3%, 1 ヶ月 16.4%であった。また企業規模が大きくなるほど、予算編成期間が長くなる傾向が見られた。Libby and Lindsay(2010)が行ったカナダ・米国の予算編成期間の調査では予算編成期間の中央値が 6 週、米国が 10 週であり、多くの企業が予算編成に 2 ヶ月以上の期間を要していることが

明らかとなった。早川ほか(2018)と清水(2016)の調査からも3ヶ月以上かけている企業が多く存在することが確認できる。

図表2 予算管理の手間・予算管理のコスト

	Q.予算の手間・コストがかかる	
	平均値	標準偏差
清水ほか(2021)	4.72	1.43
上東(2021)	4.76	1.53
清水(2016)	5点リッカート:3.48	1.04
岸田(2013)	4.15	

出典：筆者作成

図表3 予算編成にかかる期間

	上東(2021)	早川ほか(2018)	清水(2016)
1ヶ月未満	-	0.60%	2.1%
1ヶ月	16.4%	6.70%	11.1%
1.5ヶ月	-	4.50%	13.2%
2ヶ月	29.3%	16.20%	24.9%
2.5ヶ月	-	12.30%	-
3ヶ月	39.7%	32.40%	30.2%
3.5ヶ月	-	8.90%	-
3ヶ月超え	-	-	18.5%
4ヶ月	12.9%	12.30%	-
その他	1.7%	6.20%	-

出典：筆者作成

実態調査からもわかるように実務では予算管理業務の負担の減少といったニーズが生まれており、現在の予算管理システムや経営管理システムでも作業の簡便化機能や部門の予算情報伝達機能を重視した製品が多く売り出されている。しかし、横田・妹尾(2011)の

調査では「予算編成に係る時間は長期化している」の問いの平均値が 4.00 と 7 点リッカー
ト尺度の平均値であったため、以前から予算編成に掛かる時間は変化ないということにな
るが調査から 10 年経った現在ではシステムの発達により予算編成に掛かる時間が短期化
している可能性がある。

予算編成にかかる期間に関して AI による問題解決を図る場合、AI による自動化アプロ
ーチを検討することとなる。AI による予測や分析、Power BI などによるデータの可視化な
どによって間接的に業務を支援し手間を削減する可能性もあるが、予算管理プロセスの自
動化を経なくては予算管理にかかる期間の削減は行えないため、予測や分析といったアプ
ローチは他の実態調査の問題点で検討することとする。

予算編成にかかる期間の調査では半数以上の企業が 1 年のうち 3 ヶ月以上の期間を予算
管理に費やしていることが判明した。よって、管理会計担当者からは予算管理業務プロセ
スの自動化、AI 化が望まれていると解釈できるが、AI を活用しても予算管理業務プロセ
スの完全自動化には戦略面、動機付け、コミュニケーションなどの問題点が介在しており
予算管理自動化による逆機能を引き起こす可能性があり自動化を行うべきではない。

図表 4 予算管理におけるコミュニケーションの重要性

	Q.予算管理上のコミュニケーション	平均値
岸田 (2013)	企業の中で果たすべき役割について部下に伝えている	6.12
	戦略について予算数値と絡めて部下に伝えている	6.00
	目標の妥当性について部下に伝えている	5.71
	部下に現場の状況を聞いている	5.78
	部下に現場の問題点を聞いている	5.72
	市場環境について上司に積極的に伝えている	5.82
	問題点を上司に積極的に伝えている	5.89
	部下はトップマネジメントから何を期待されているかを理解し ている	5.49

出典：筆者作成

予算管理業務を行う上で、トップ層とミドル層やローワー層、もしくは親会社と子会社間

でコミュニケーションを行うことは重要である。参加型予算編成プロセスでは従業員と経営者層の交渉が重要であり、繰り返し交渉を繰り返し合意の上最終的な資源配分が決定される。AIによる完全自動化を行った際には予算管理プロセスにおけるコミュニケーションで生じるメリットが失われてしまう可能性がある。岸田(2013)の予算管理に関するコミュニケーションに関する調査票からもコミュニケーションは大切にされており、妹尾(2017)の質問票調査においても「当初の予算目標と実績が乖離した場合、事業単位の上層部とミドルマネージャーの話し合いがもたれる」の平均値 5.32, 「事業単位の上層部は予算実施経過の報告を定期的を受け、ミドルマネージャーとの話し合いが定期的にもたれる」の平均値 5.34 と予算管理業務においてコミュニケーションを重視している企業は多いことがわかる。

安達(2017)では IT 企業である現場サポートにおける予算管理と人的コミュニケーションを調査した。現場サポートでは予算達成のための環境変化へ迅速な対応手段として組織成員間の情報共有、共通認識の醸成を頻繁に行っている。飛田(2012)では中小企業を対象として規模間別に組織成員間のコミュニケーションを調査した。人間関係は業務を行う上で重要であり、企業の規模に関係なく重要なものであるが、その傾向は face to face の小規模組織の方が顕著である。規模が拡大していくにつれて経営管理や会計システムの重要性が高まるのかもしれないと論じている。鬼塚(2018)では本社が業績管理システムを利用することで在外子会社の意思決定に与える影響を調査した。本社と子会社が頻繁にコミュニケーションを取ることにより、認識ギャップを埋めることができ、本社の意思決定を子会社に反映することができることが示唆されている。

谷守(2019b, 90)によれば「予算のとりまとめでは、各部署間でのコミュニケーションが重要である。たとえば、ボトムアップ予算は、現場から要求される予算が本部へ報告され、本部ではそれをとりまとめた後に総合予算が作成されるが、実際には予算スラックの問題から、売り上げ予算については低めになり、原価予算は高めになる。そのため、本部では現場の各部署との調整(いわゆるストレッチ)が重要になる。しかし、いまのところ AI では本部と現場との微妙な思いの絡む調整など到底できない。」と主張している。日本企業は稟議をもって予算管理の実施を行っており(伊藤 2007)、予算管理業務プロセスに稟議制度を導入することで情報共有機能や責任権限の明確化を図ることができる(上山 2021)。よって、業務の完全自動化をおこなった場合、情報共有機能や責任機能が失われる可能性がある。Dahlanm(2019)は公共サービス行政におけるミドル、ローレベル層の従

業員における管理会計システムと対話型予算の利用の重要性について調査を行った。その結果、管理会計システムの利用によってトップレベルの従業員がミドル、ローレベルの従業員に対してコミュニケーションチャンネルを構築し、仕事のパフォーマンスを向上させることができるということが判明した。

これらの先行研究から予算管理プロセスにおけるコミュニケーションを通じて、組織成員間の情報共有、共通認識の醸成、認識ギャップの防止、予算スラックの調整といった機能を実現させていることが判明し、予算管理自動化 AI によってこれらの機能が消失してしまう可能性を有していることから予算管理の完全自動化 AI は現場に不利益を被る可能性が想定される。

また、AI を導入することで、従来の予算管理よりコストが増える可能性がある (Marotta and Au 2021) という指摘もあり十分な予算管理の質を担保しつつ、予算管理にかかるコストを削減することも容易ではないといえるだろう。AI を導入する上では従来のコンピューターシステムに比べ、多くのイニシャルコスト、ランニングコストを要することとなる。以前までのシステムであればパッケージやソフトウェアを導入すれば十分に効果が発揮できていたが、AI システムは運用が必要となるため、データの運用費や専門家の人件費といったランニングコストが生じる。将来的にさまざまな企業形態に対応できる完全自動化予算管理システムの実現によって、予算管理の手間やコストが削減されることは考えられるが、その場合においても上記のコミュニケーションや業績管理といった予算特有の問題が発生する可能性が高く、予算管理の AI による自動化は困難 (谷守 2021) という先行研究を否定することは難しいといえる。

4.3 予算管理の環境変化に関する実態調査レビュー

Hope and fraser (2003) によれば予算は期首に編成された計画を前提としており、事業環境の不確実性の高まった近年では予算の前提変化があり得るとしている。予算は予算編成の頻度を高め迅速な対応を行うアプローチとして予算編成のタイミングを細分化し、一定の期間経過時ごとに当初予算期間の予算を再編成するローリング予算や期中に必要な応じて予算を修正し、実状にあった予算へと頻繁策定するといった考え方もある。

図表5 予算管理の予算の期中変更

	Q.当初予算の変更は行わない		
榊谷ほか (2022)	年度	平均値	標準 偏差
	2009年	3.19	1.76
	2014年	3.50	1.78
	2019年	3.63	1.99
早川ほか (2018)	Q.予算の変更		
	予算修正なし	17.9%	
	必要に応じて修正	68.2%	
	ローリング予算	14%	
清水(2021)	Q.短期の予測に基づいて当初予算が公式に変更される		
	変更される	50%	
	変更されない	50%	
清水(2021)	Q.短期の予測に基づいて当初予算が公式に変更することは無いが、組織内で着地点を新たに設ける		
	着地点を設ける	33.7%	
	着地点を設けない	19.1%	
	未回答	47.2%	

出典：筆者作成

予算の期中に見直しを行う企業を調査した実態調査では、早川ほか(2018)68.2%、清水(2021)50%と基本的には半数の企業が予算の期中変更が行われている。榊谷ほか(2022)では「状況変化にかかわらず、当初予算目標の変更はしない」という調査を10年間にわたり実施しており、当初予算の点検・修正を行わない企業が増加しているといった調査結果がでていいる。また清水(2021)では特定の状況下でどう予算の変動を行うかについての調査も実施しており、売上予算の変更について、「業界全体の市況が良く、競合他社の業績も良い場合、予測値が予算を上回る時、売上高予算は上方修正される」という問いに対して、平

均値 5.27 と高い結果が出ており、市況が良い場合上昇志向が強く出ることが判明した。同じ状況下で費用予算の変更についての調査は平均値 4.11 であり、環境が良くても予算の上方修正を行うわけではないという結果となった。逆に環境が悪い場合での調査、「業界全体の市況が悪く、競合他社の業績も悪い場合、予測値が予算を下回るとき、売上予算は下方修正される」かの問いについては、平均 5.27 であり、環境が良い場合と逆の対応を取ることが判明した。費用予算については平均値 5.11 となっており、環境が良い場合と比べて、保守的な傾向が見られた。

図表 6 予算管理上の環境変化への対応

	Q.環境に対応することができない	
	平均値	標準偏差
清水ほか (2021)	3.66	1.46
上東 (2021)	4.76	1.31
清水(2016)	5点リッカートン:2.74	0.88
横田・妹尾(2011)	4.85	

出典：筆者作成

次に企業が環境予測の対応についての実態調査を整理していく横田・妹尾(2011)の調査では「予算編成時における利益の予測が難しくなった」という問いに対して平均値 5.37、上東(2021)では「環境予測の困難性」の平均値 4.85、清水ほか(2021)「環境変化が激しく、予算はこれに対応することができない」平均値 3.66 であった。筆者は新型コロナウイルスの影響によって環境の予測が行いづらく対応に苦労している企業が増加したと予想していたが、そのような結果は質問票調査では見受けられなかった。考えられる事象としては「テクノロジーの進歩により、予測の精度が純粋に向上した」「新型コロナウイルスの影響により環境予測に対する投資を行い結果として対応力が向上した」「新型コロナウイルスの影響により環境予測に対する意識が以前よりも低下(予測を諦めた)した」などがあげられるだろう。

期中に予算の見直しを行う企業のみを対象とした環境予測の困難性の調査を実施した先行研究は筆者の知る限りでは存在せず、予算の期中見直しを行う企業はその他の企業と

比べ、予算に対する意識が高いと想定され、そのような企業が環境予測に苦しんでいるのかどうかを調査することは意義があるといえる。

期中に予算を変更している企業は 5~6 割強であり、年代による変動はあまり見られなかった。しかしながら、予算の期中の予算変更は行わないもののロードマップを用いた目標の変更を行うケース(堀井 2015)も存在するため、予算に対して期中に何かしらのアプローチを行っている企業は多いことがわかる。AI による予算管理では多量のデータやビッグデータを取り扱う関係上、リアルタイムな環境変化に対応し、期中の予算を適切に修正するアプローチが期待されるが、期中に予算の変更をおこなっていない企業は行っている企業に比べ、予算に関する関心が薄いことや期中に変更された予算に対応ができないことが想定され、管理会計の AI システムの利活用において遅れを取ることが考えられる。

実態調査では環境予測への対応を問題としている企業が以前と比較して減っていることが明らかとなったが、前述した通りイレギュラーな事態により予測自体に意味がないとし、精度の高い予測を諦めた企業がいることが推測される。新型コロナウイルスなどの事態の予測は AI で行うことはできないが、経済環境の予測、他者動向の予測、経済市場の予測、政策等による影響の予測など予測タスクに関することは AI の得意とする領域である。人間が扱っている情報量と比べ、圧倒的な量の情報を処理できるため LSTM などの時系列予測モデルやランダムフォレストや勾配ブースティングといったアンサンブル学習モデル、Transformer による長期依存関係を把握するモデル、DQN などの強化学習モデルなど様々な予測モデリング手法が考えられる。しかし、Losbichler and Lehner(2021)が指摘しているように AI モデルでは経済事象全体のモデリングは困難であり、完全な予測を提供することは難しいことから利用者は予測情報を適切に処理し、自社の状況を鑑みた上で予算管理に役立てる必要がある。

4.4 目標水準、戦略とのリンクに関する実態調査レビュー

予算編成の目標と事業戦略がリンクしているかの実態調査では、梶谷ほか(2022)が 2009 年、2014 年、2019 年に行った「重視する業績目標は事業戦略と整合性が取れている」という問いに対して、平均値の推移 4.84→5.01→5.26 と年々事業戦略と業績指標の整合性が高まっていることを明らかとした。山田ほか(2003)によれば 1992 年と比べて、2002 年は予算と中・長期経営計画や短期経営計画とのリンクが強化されてきているという調査もあ

るため、予算の設定と戦略の意識が昔よりも高まっていることが窺える。

脱予算論者によれば予算は戦略的に重視されることはほとんどなく、矛盾していることが多いと主張されていた(Kaplan and Norton 1996;Neely et al. 2003)が、現在ではその問題は解決に向かっているといえるだろう。予算自体の設定の調査でも妹尾(2017)では「予算や業務に関する手順・手続きは明確である」5.31,「予算計画はきめ細かく設定されている」5.34 と予算編成プロセスに力を入れている企業が多く予算設定への意識の高さが窺える。

図表7 予算と戦略のリンク

	Q.業績目標は事業戦略との整合性が取れている		
	年度	平均値	標準偏差
梶谷ほか(2022)	2009年	4.83	1.23
	2014年	5.01	1.31
	2019年	5.26	1.30
妹尾(2017)	Q.戦略的な予算編成が行われている	3.82	1.13
岸田(2013)	Q.予算目標の達成が戦略と結びついていない	3.09	1.65

出典：筆者作成

予算管理において戦略面とのリンクが強まってきており、戦略を反映させた予算管理を行うことが求められていることが判明したが、AIにその戦略をどう考慮させるかがAI活用の観点からは問題となる。特に短期的な戦略は変更が生じることが多くあり、企業の戦略方針の変更が生じた際に柔軟性に欠けるAIモデルでは対応できずにモデルを組み直すことを要する事態になりかねないため、注意が必要である。

企業戦略への対応を考慮すると様々な要求に応えられるモデルが望ましいため、管理会計タスクに特化した自然言語処理によって柔軟に対応するモデルであるLLMが想定される。また、予算管理タスクを自然言語でAI支援を行う上では文脈を考慮した上での支援の実現により一貫性を持ったAIの応答が求められるため、長期的な文脈構造の忘却を避けるように計算コストの工夫がされたモデル(Zang et al. 2021;Dai et al. 2020)による構築がされることが要求されるだろう。

4.5 予算の設定に関する実態調査レビュー

清水ほか(2019)によれば先行研究上の予算の目標値は暗黙の前提として売上高・費用・利益の目標値を想定しているとされている。しかし、上東(2021)の調査では「損益予算、資金予算、資本予算、販売予算、製造予算、研究開発予算」といった各種予算についての重視度合いについて調査を行った。「損益予算」と「販売予算」が平均値6を超えており、非常に重要視していることがわかった。「資金予算」、「製造予算」については企業の予算の重要度では中間といった具合であり、「研究開発予算」は大規模企業になるほど重要視する傾向にあった。対象となった予算の中で一番重要視されていない予算は資本予算という結果となった。この調査結果となった要因として考えられるのは、資本予算は数年間にわたる投資に関わる予算であるため、編成時に調整されることはあっても大きな変更を伴うことは頻繁に起こり得ず、検討の機会が少ないことから、このような結果となったことが推測される。

予算編成プロセスでは管理職が主導的な役割を果たすか、現場の意見が主導的な役割を果たすかの区別をすることができる。理論・実務においては「トップダウン」「ボトムアップ」「対話型アプローチ」に区別して目標値の設定を行う(Marotta et al.2022)とされるが、上東(2021)では予算方針の策定についてより細かく項目を分け「予算事務が原案を作成し、トップが承認」「トップが方針を提示し、予算事務が具体案を作成」「上位部門が原案を提示し、予算事務が調整の後、トップが承認」「トップが具体方針を提示し予算部門で調整」「その他」の5項目の調査を行った。大規模企業においては「上位部門が原案を提示し、予算事務が調整の後、トップが承認」といった形式をとっている企業が一番多く、42.5%であった。次点では「予算事務が原案を作成し、トップが承認」の30.1%であり、「トップが方針を提示し、予算事務が具体案を作成」は23.9%、「トップが具体方針を提示し予算部門で調整」は3.5%とほとんど存在せず、その他は0%という結果であった。

図表 8 予算の決定プロセス

清水ほか (2021)	Q.最終的な予算はトップが 決定している	平均値	標準偏差
		5.73	1.25
上東 (2021)	Q.予算方針の決定 (大規模企業)	割合	
	予算事務が原案を作成し トップが承認	30.10%	
	トップが方針を提示し予算事務が 具体案を作成	23.90%	
	上位部門が原案を提示し予算事務が 調整の後トップが承認	42.50%	
	トップが方針を提示し 予算部門で調整	3.50%	
妹尾(2017)	Q.対話型の予算編成を行なっている	平均値	
		5.2	1.13

出典：筆者作成

上東(2021)の調査で企業が重視している予算の調査を行ったが、AI のモデルとして LLM を仮定した場合、それぞれの予算について別々の予測を行うといったことは行わない。LLM であれば管理会計担当者が管理したい予算についてプロンプトを入力することであらゆる予算に対応できるためである。しかし、学習に用いるデータの優先度を考慮する際には調査に基づいて関心の高い予算のデータを優先的に収集・加工すべきである。

予算の決定プロセスでの実態調査では、予算の最終決定はトップが行う企業が多いという結果となった。管理会計担当者が AI システムを用いて予算の策定を行った場合、それをトップが十分に理解できる根拠を持った説明が求められる。しかし、谷守(2020)でも指摘されているように AI はブラックボックス問題を抱えており、最先端の AI であれば以前までの AI よりも計算過程が複雑化しており、より顕著に現れることとなる。LLM を用いた予算編成 AI であれば予算管理プロセスにおいて、AI と管理会計担当者が対話型のやり取りを行うこととなるため、その対話履歴がある種の説明根拠となる可能性もある。

ブラックボックス化の対応として Explainable AI の開発が進んでいるが、Transformer 派生のモデルは非常に複雑であることから算出した答えが可視化されても人が理解することは難しいと思われる(LLM Visualization といった可視化モデルは出てきてはいるが説明根拠とするには厳しい)。LLM 自体にプロンプトでなぜその予測を行なったのかを聞くことは可能であるが、内部の処理まで踏まえた説明をしてくれるとは限らず、計算過程も把握した上で AI による支援を求める場合 Transformer モデルでは限界があるため、計算理論が体系的に構成されているモデル(Hippo 系モデル(Fu et al.2022)など)を用いたアプローチを検討することでブラックボックス問題を解決できる可能性がある。

4.6 予算スラック・ストレッチな予算に関する実態調査レビュー

予算スラックの調査では平均的な回答が見られ、ストレッチな予算を設定しているかという調査では榎谷ほか(2022)によればストレッチな予算をとっている企業は増加傾向にあるものの、案外ストレッチな予算を設定している企業は多くないという結果となった。

「前年度目標を上回るように予算を設定している」かの項目では清水(2016)では2割の企業が前年度ベースでの予算を設定していることが判明した。これらの実態調査では他の調査に比べ標準偏差が高い傾向にあり、企業ごとの違いが現れた結果、平均的な数値となったと解釈できる。

図表9 予算スラックの形成

	Q.予算の目標値が低くなりがち(予算スラック)	
	平均値	標準偏差
清水ほか (2021)	3.48	1.54
上東 (2021)	4.04	1.32
清水(2016)	5点リッカート: 2.68	0.97

出典：筆者作成

図表 10 予算のストレッチ性

	Q.ストレッチな予算をとっている		
	年度	平均値	標準偏差
榎谷ほか(2022)	2009年	3.59	1.03
	2014年	3.56	1.99
	2019年	3.89	1.58
妹尾(2017)	-	3.55	1.39

出典：筆者作成

図表 11 予算目標の設定

	Q.前年度目標を上回るように設定している	
	平均値	標準偏差
清水ほか (2021)	4.42	1.61
清水(2016)	20.1%	

出典：筆者作成

従来、予算編成において目指すべき目標値はストレッチな予算であるべきという主張がなされてきた。しかしながら、本当にストレッチな予算が従業員のモチベーション、パフォーマンスひいては企業の業績向上に繋げることができるのであろうか。ストレッチな予算が企業業績に与える影響に関しては明らかとなっておらず、あくまでストレッチを良いとする見解が多いに過ぎない。実際の企業では予算スラックが組み込まれているケースが多く、予算スラックは適切に利用することで組織にバッファーを持たせ、長期的な貢献に寄与している可能性もある(岸田 2020)。AI において目指すべき予算とは、究極的に業績の向上が図れる予算である。業績の向上に起因する事象として大量のデータのリアルタイム処理による環境変化への対応や従来多くの工数を用いる予算編成の効率化によるコスト削減、分析の高度化による予算統制の精度向上などが挙げられるが、これらの AI による利点をどのように活用し、どのような予算の値を決定するのかを探索していく必要がある。現時点においてはできるだけストレッチな予算に近づけることが無難な選択であり、AI で

予算管理を行う際にもひとまずはストレッチな予算を算出することになるだろう。しかし、将来予算に関する多種多様な企業のデータが学習された後の AI はストレッチな予算とは異なる予算を算出してくることが考えられる。目指すべき目標値を定量的に示すことは、従来の予算管理プロセスでは困難であり現実的ではなかった。AI に予算を考えさせ、算出される解を分析することによって目指すべき予算像に変革が起こる可能性もあるのではないか。

4.7 予算の目的に関する実態調査レビュー

企業にとって予算とは企業の利益目標を達成するための経営管理プロセスであり戦略と予算がリンクしていることが望ましい。予算は短期の計画に過ぎないが、日々の企業の業務活動に合わせて編成されるため、業務上の指針となるものである(伊藤, 目時 2021)。これらの予算管理に対する考え方は実務において学術的な予算の目的と整合しているのだろうか。

妹尾・横田(2013)が行った予算の目的に対する効果への影響を調査した結果から平均値が 5 点以上の目的を抽出すると「最終目標や方針の伝達」「達成すべき重要事項の周知徹底」「目標の具体化」「マネージャーや従業員の動機付け」「迅速な改善措置の促進」「責任の割り当て」「組織単位への資源配分」「組織単位間の調整」が挙げられ、動機付けや目標の明確化といった項目が重視されていることがわかる。

上東(2021)の調査では上記以外の項目として「所要の収益性の実現」や「部門の業績評価」が予算管理の目的として強調されていた。加えて企業規模が大きくなるほど予算に期待する効果が大きくなることが判明した。「最終目標や方針の伝達」「達成すべき重要事項の周知徹底」といった項目は企業の戦略に結びつくものであり強く意識されていることがわかる。他にも動機付け、責任、資源配分といった活動が予算管理によって支援されていることがわかる。

予算管理には複数の型があるとされ、予算の特徴に応じて 5 つに分類可能である。Sponen and Lambert(2016)は予算管理の特徴を変数としたクラスター分析により「ヤードスティック予算」「強制的予算」「インタラクティブ予算」「ルース予算」「目安予算」に分類した。報酬とのリンク性が高く予算編成においてマネージャーが参加し、予算プロセスにおけるシニア・マネジメントの関与があり、目標の難易度が低く予算修正の頻度はそこ

そこであるインタラクティブ予算が最も予算への満足度が高いことを明らかとなった。妹尾(2017)ではその分類を日本的な予算管理に置き換えて類型化し、クラスター分析により3つのクラスターに分類した。挑戦的でなく達成可能な業績目標を維持し予算目標の弾力的変更を行う「弾力的予算編成」(39.3%), 予算の変更をあまり行わず事前, 期中, 事後の予算管理プロセス全体の統制を重視する「統制型予算編成」(40.6%), 積み上げ型の予算を採用しており, 予算管理プロセスへの関心が薄い「非統制的予算管理」(20.1%)に分けられると論じた。

図表 12 予算管理の目的

	Q.予算管理の目的(大規模企業)	平均値	標準偏差
上東 (2021)	収益の実現	6.32	
	財務安全性の確保	4.89	
	原価の引き下げ	5.00	
	部門の業績評価	5.81	
	資源配分の有効性の達成	4.33	
妹尾・横田(2013)	最終目標や方針の伝達	5.81	0.92
	達成すべき重要事項の周知徹底	5.69	1.00
	目標の具体化	5.88	0.90
	マネージャーや従業員の動機付け	5.32	1.03
	迅速な改善措置の促進	5.01	1.19
	責任の割り当て	5.48	1.12
	組織単位の資源配分	5.33	1.02
	業務量の決定	4.51	1.23
	組織単位の調整	5.17	1.04

出典：筆者作成

予算に求められている役割として、動機付けや伝達といったコミュニケーション要素、資源配分や収益性実現といった会計的要素、責任の割り当てや組織単位間の調整といった組織内関係の要素が求められていることが明らかとなった。予算編成 AI システムではこ

これらの要素の支援、もしくはこれらの要素を無くさないことが求められる。例えば資源配分や収益性実現はリソースの最適化を目指すモデルや環境の予測によって支援できる可能性がある。コミュニケーション要素に関しては予算管理にかかる期間やコストの実態調査レビューでも論じたようにシステムによってトップ層が現場を見る機会を失うことは避けなくてはならない。むしろ、Dahlanm(2019)や Sponen and Lambert(2016)にもあるように予算管理のコミュニケーションを通じて組織のパフォーマンスが向上するのであれば、システムを通じてコミュニケーションを増加させる機会ができるよう予算管理行動を調整することが望ましい。組織内の責任の割り当てや調整といった目的は AI システムによって解決を図るのではなく、目的を阻害することのないようにシステムを構築すれば十分である。

4.8 業績評価に関する実態調査レビュー

妹尾・横田(2013)が行った予算の目的に対する効果への影響の調査項目について探索的因子分析を行った結果では、「予算による垂直的コミュニケーション機能の向上」「予算による計画機能の向上」という因子が抽出された。予算に基づく業績評価では主観的評価を行う仕組みがあると評価結果とボーナスのリンクの強化によってコミュニケーションが活発になると論じた。また主観的評価は予算による計画機能に負の影響があることが判明した。梶谷ほか(2022)によれば以前よりも業績管理において財務指標重要度、業績目標水準(挑戦的な目標水準)が高まっていることが指摘されている。また同調査にて「個人・部門業績の評価は、業務プロセスより結果が問われる」という問いに対し、平均値が以前に比べ減少傾向にあり、業績評価においてプロセスも重視されるようになってきている。しかしながら、平均値が4点を超えており、結果が重視されていないわけではないことに留意しなくてはならない。また、業績単位の長やミドルマネージャーのレベルにおいて業績報酬のリンクが強い企業では結果、弱い企業ではプロセスを重視する傾向が示唆できると主張している。清水(2021)の調査にて「事業部・部門長の業績は、当初予算の達成度によって評価される」の問いでは平均点 5.35 と企業が業績評価において当初予算をいかに重要視していることがわかる。

図表 12 予算と業績評価

		平均値	標準偏差
清水ほか (2021)	Q.業績は当初予算の達成度で評価される	5.35	1.26
清水(2016)	Q.予算の達成度は賞与に反映されない		
	反映されない	17.1%	
梶谷ほか(2022)	Q.個人、部門の業績評価は業務プロセスよりも結果が問われる		
	年度		
	2009 年	4.53	1.03
	2014 年	4.05	1.15
	2019 年	4.01	1.15

出典：筆者作成

本論文における管理会計 AI の構想では業績評価について度外視している。業績評価は予算管理プロセスの一端を担う重要な要素であるが、AI の構築においては予算編成 AI とは全く別のアプローチとなることが考えられるため、この項での検討に留めることとする。業績評価に関する実態調査の先行研究では業績報酬との関連性の強い企業では結果、弱い企業ではプロセスが重視されることや当初予算の達成によって事業部・部門長の業績が強く評価されること、予算目標の達成難易度に関わらず予算が達成されたか否かによって評価が行われる傾向にある(清水ほか 2021)ことが明らかとなった。(清水 2021;梶谷ほか 2022)結果を重視するかプロセスを重視するかという点は企業の指針によるものであり、その企業のポリシーに準じて決める要素であるため、AI の介入余地はないといえる。しかし、「当初予算が強く意識される」ことは一部企業の指針的要素も含まれるものの予算が環境変化に対応できていない可能性も示唆される。さらに「予算目標の達成難易度に関わらず予算が達成されたか否かによって評価が行われる傾向」は評価対象である部門や部門責任者からすると公平性に欠けるものである。この 2 点においては AI の適用により問題点緩和の可能性を秘めている。

1. 「当初予算が強く意識される」

当該問題は本論文で構想を行なっているような予算編成 AI を導入することで解決される可能性がある。予算編成 AI では大量の環境予測要因となる情報をリアルタイムに処理し予測に役立てることができる。結果、随時更新される予算情報を活用することにより当初予算の持つ意味合いは薄れていき、期中の状況を勘案した業績評価につなげることができるだろう。

2. 「予算目標の達成難易度に関わらず予算が達成されたか否かによって評価が行われる傾向」

当該問題は予算の達成目標の把握が適切に行えないことによって生じる問題である。よって AI による予算目標の評価モデルを検討することとなる。予算に対して難易度を判別するとなると AI による事後的な評価が有効である。それぞれの部門が持つ予算と経済環境、市場動向、他社動向、過去の予算実績を参照し事後的に予算の達成難易度評価を数値化し業績評価値の重みとすることで達成難易度を反映した業績評価が行えるだろう。

もしくは事前アプローチとして、部門が予算の最終調整を行う企業の場合であれば部門が提示してきた予算の難易度を評価することで、部門の予算設定時の目標値をよりタイトなものへと促せる可能性もある。難易度数値化プロセスにはランダムフォレストや勾配ブースティングなどの決定木モデルを使うことが考えられる。

安達ほか(2022)は人事評価のフィードバックについて AI と人間どちらが公正感や納得感を持つことができるのかについて調査を行った。

図表 13 人間と AI による評価の感じ方の違い

	受け取った評価の納得感	受け取った評価の公正感	フィードバック者(上司 or AI)への信頼
人間	4.305	4.199	3.199
AI	4.12	3.699	2.709

出典：安達ほか(2022)をもとに筆者作成

人間と AI が示すフィードバックは同様の内容である。受け取った評価の納得感、受け取った評価の公正感、フィードバック者(上司 or AI)への信頼いずれにしても現時点では人間の方が上回っているが大きく乖離はないといえるだろう。この調査は 2022 年 6 月に公開された文献によるものであるが、調査時期が不明であったため、調査時点において

ChatGPT や GPT4 が公開されていたかは定かではないが、生成 AI の普及により信頼性が人間を上回る可能性も考えられる。

4.9 予算実績差異分析に関する実態調査レビュー

予算管理プロセスは予算編成と予算統制に分けることができる。予算統制の重要な要素として事前に設定された予測の数値と事後的に把握される実際の数値を比較し、差異を導きだし将来的な計画に役立てるために予算実績差異分析を行う。予算実績差異の把握を行う目的としては改善措置としての手段であり、コントロールを適切に行う上では必要不可欠なプロセスである(企業予算制度研究会 2018)。

中村(2015)では予算の中で問題視されている機能は統制機能に関わるものが多いが、成果主義導入の失敗例が多いことから統制機能に関わる予算の問題は日本の企業ではそれほど気にしなくて良く、予算の調整機能の重要性を考慮すると欧米に比べ日本の予算有用性はいまだに高いと考察している。上東(2021)では予算実績差異分析の目的に関する調査を実施し、「改善措置」「部門成果の評価」「差異の報告」に強い関心があることが判明した。

2022 年に行なった予算の状況について研究室の調査⁴では予算実績差異を把握している企業は 91.8%であった。把握している差異の内訳として、一番多く把握されていたのは「予算総額と実績総額の差異」であり、予算実績差異を把握している企業のうち 71.3%が実施していた。次点では「費目ごとの予算と実績の差異」の把握を 57.5%の企業が実施している結果となったが、他の差異の把握を行なっている割合は半数以下であった。また予算実績差異の報告について「予算実績差異について、差異が生じた原因分析についても報告を行う(報告を受ける)」という問いに対して 5 点評価中平均点 3.50、「予算実績差異分析の結果はタイムリーに関係者間で共有されている」に対して平均点 3.41、「予算実績差異分析の結果を受けて、改善のための対応策を考えている」に対して平均点 3.56 と予算実績差異の報告共有を重要視している企業が多いことがわかる。予算実績差異報告の頻度について(複数回答可)それぞれの差異ごとに調査しており、平均してみると月次→半期→四半期→年度→週次→日次となった。中でも月次ベースで差異を把握している企業は 7 割を超えており、月ごとに差異を把握している企業が多いことが窺える。

⁴ 古賀絢子.2022.「予算管理におけるフィードバックとフィードフォワードの関係」『Institutional Repositories DataBase』:1-38.

図表 14 予算実績差異分析の目的

		Q.予算実績差異分析の目的	
上東 (2021)	改善措置		5.63
	部門成果の評価		5.25
	差異の報告		5.73
	部門主管者の評価		4.68

出典：筆者作成

図表 15 予算実績差異分析で把握する差異

	予算実績差異を把握している企業	n=167	割合
1	予算総額と実績総額の差異	119	71.3%
2	費目ごとの予算と実績の差異	96	57.5%
3	直接費の価格差異	58	34.7%
4	直接費の数量差異	52	31.1%
5	セールスマックス差異	38	22.8%
6	市場占有率差異	36	21.6%
7	製造間接費予算差異	58	34.7%
8	製造間接費能率差異	43	25.7%
9	製造間接費操業度差異	31	18.6%
10	その他【 】	1	0.6%

出典：研究室の調査(2022)をもとに筆者作成

予算管理 LLM を活用することで多くの企業が統制活動として実施している「予算総額と実績総額の差異」の差異要因を多くのデータに基づいて自然言語ベースで解析することが可能となる可能性を秘めている。また、時系列予測の機械学習モデルを利用することで過去のデータから学習し、トレンドや季節性を捉えられる可能性もある。

Casas-Arce et al.(2017)によれば、会計情報のフィードバックは頻度が低い時に最良のパフォーマンスを出すことが明らかとなっている。中村(2015)での指摘通り、日本と欧米では状況が異なるため、Casas-Arce et al.(2017)の研究が日本でも適用されるかは未知数であるが、AIによってリアルタイムに環境状況を踏まえたフィードバックを行える状況が整ったとしても、フィードバックの頻度が高すぎた場合業績がむしろ低下する可能性を示唆している。

第5章 予算編成 AI の構造と構成

執筆時点において管理会計における AI の適用は学術的にも実務的にも十分な検討がされているとは言い難い Ulrich et al.(2022)。学術的な視点では AI と会計について論じる際に具体的な技術面に触れているケースは少ない。機械学習や深層学習といった区別や教師あり学習，教師なし学習，強化学習の区別がなされているケースは存在するものの具体的なアルゴリズムの追求をしているケースはかなり限られる。また先行研究では AI によって可能なことや不可能なことについて論じていることが主流であるが，AI の技術進歩は非常に早い速度で進んでおり，以前では不可能とされていたことが現在においては可能であると考えられる点も珍しくない(谷守 2023)。特に 2022 年に公開された ChatGPT は大きな衝撃を与え，世の中の AI に対するイメージを大きく変えた。谷守(2019a)が今現在の AI では創造性，芸術性，戦略性，フィロソフィーなど感情が関係する領域の業務やコミュニケーションを要する業務は困難だと論じていたが，現在では生成系 AI によってコミュニケーションへの対応はある程度実施できる。感情についてもテキスト，画像，音声から感情分析は可能なため対応できると推測する。

依然，管理会計に AI を適用するとなると「教師データ不足による経済のモデリングの不十分性」「ブラックボックス」「勾配消失」「長いシーケンスへの対応」といった課題は有しているが具体的な AI モデルの構想を練るには十分な技術進歩がなされた。AI のコミュニケーション能力向上と信頼の確保が現実的になった今日では，あらゆる管理会計論点へ管理会計の適用可能性が向上したことは間違いない。

5.1 予算編成 AI の構想

本論文では予算管理に対し AI 適用を検討した。予算とは会計数値を用いて企業の戦略を表現するものであり，企業の業務活動へ反映させるための指針である。予算管理を軸とした教師データを学習させた AI モデルを構築することで他の管理会計にも準用できると考えられ，そのようなシステム構築を行うことで資源の無駄なく活用に貢献することが期

待される。先行研究レビューの結果では企業が予算管理に抱えている問題の一つが手間・コストの問題であることが明らかとなったが、AIによって当該問題の解決を図った場合、完全自動化もしくは多くの部分を自動化するというアプローチが考えられる。しかし、予算管理の完全自動化には複数の逆機能が存在していることから単に予算管理を自動化するシステムを構築すべきではないといえる。

---予算管理完全自動化の逆機能---

- ・ 予算管理プロセス上のトップ層とロー層のコミュニケーション機会の喪失によるトップ層の現場の理解が損なわれてしまう可能性
- ・ 企業戦略の反映が困難
- ・ 個人、部署、事業部の能力の考慮が難しく適切な動機付けを行うことが困難
- ・ AI 予算のブラックボックス問題

完全自動化を行った際にはこれらの逆機能が生じる可能性があり、予算管理プロセスの機能を十分に発揮することが難しい。今後、管理会計領域での AI 運用が普及し、自動で最適な予算を策定できるレベルのデータが集まればこれらの逆機能を補うほどの成果が出る可能性もあるといえるが、現時点で行える自動化では逆機能を避けることはできない。

実態調査において、最も予算管理の課題とされていた項目は、予算管理にかかる期間やコストの問題であったが、管理会計部門の状況、課題を明らかにするために研究室⁵で行った質問票実態調査では異なる結果となった。予算管理システムを構築する上で必要と思わしき要素を顕在化する目的で「予算管理システムについての課題として認識されていること」について自由記述形式で調査を行い、その要素を集計したのが図表 16 である。

⁵ 詳細は工藤桂菜.2024.「高業績企業に共通する管理会計組織の特徴に関する研究:質問票調査に基づいて」を参照されたい。

図表 16 予算管理システムの課題

Q16. 予算管理システムについての課題として認識されていることがあればご記入ください。（要素集計時に重複あり）		
1	システム化	12
2	作業上の負担	1
3	精度の問題	3
4	経理上の問題	7
5	分析上の問題	6
6	属人化	3
7	環境変化への対応	3
8	効率性の問題	4
9	その他	2
10	人材問題（属人化を除く）	3
0	問題なし	3

出典：研究室の調査(2024)をもとに筆者作成

予算管理システムの課題として、エクセル管理のみで予算管理プロセスを処理していることからシステム化を進めたいといった「システム導入・設計の課題」やグループ内での統一科目が利用できていないなどの「経理上の課題」が多く寄せられた。こういった課題を抱えている企業は予算管理に AI を導入しシステムの高度化がなされた場合、十分にその恩恵を受けることは難しい。なぜなら、AI は以前までのシステムとは異なり導入し手順を覚えるだけでは AI システムの運用はできない。追加的にデータの学習をさせ、将来的な精度の向上を目指していかなければ AI システム導入の恩恵は十分に得られず、システム面や経理面に課題を持つ企業に AI 適用を行った場合、現場は混乱を招く可能性がある。

次点で多く寄せられた意見として、予算統制時の要因探求などの「分析上の問題」やシステムにかかる工数が多く無駄が多い「効率性の問題」、純粋な予算の精度の悪さなどの

「予算精度の問題」などがあつた。これらの問題点は十分に AI でアプローチすることが考えられる。分析は LLM による解釈の支援や RNN や CNN, SVM を利用した非線形関係のモデリングなどの活用が考えられ、効率性や予算精度の問題は分析や LSTM などの時系列予測や自然言語処理を利用した環境の予測などによって改善を検討することができる。

図表 17 予算管理に用いているシステム

Q.予算編成にどのようなシステムを使用しているか	
表計算ソフト(Excel など)のみを使用	60 社(50.0%)
予算管理アプリケーションも併用	57 社(47.5%)
システムを利用していない	3 社(2.5%)
Q.着地予測にどのようなシステムを使用しているか	
表計算ソフト(Excel など)のみを使用	76 社(63.3%)
予算管理アプリケーションも併用	40 社(33.3%)
システムを利用していない	4 社(3.3%)
Q.予算の調整にどのようなシステムを使用しているか	
表計算ソフト(Excel など)のみを使用	77 社(64.2%)
予算管理アプリケーションも併用	37 社(31.0%)
システムを利用していない	6 社(5.0%)
Q.予実管理にどのようなシステムを使用しているか	
表計算ソフト(Excel など)のみを使用	46 社(38.3%)
予算管理アプリケーションも併用	72 社(60.0%)
システムを利用していない	2 社(1.7%)

出典：研究室の調査(2024)をもとに筆者作成

図表 17 は予算管理のプロセスにおいてどのようなシステムを利用しているかを調査した。予算管理アプリケーションが最も使われているのは予実管理となり、予算統制プロセスにおける分析のための可視化やフィードバックを重視していることがわかる。

図表 18 予算管理の AI 利用予定

Q.予算プロセスにおいて AI をどのように利用または利用予定ですか。	
1	妥当性を検討中
2	過去実績等からの傾向予測に AI を利用予定
3	短期, 中長期, プロジェクト損益予想
4	市場環境を踏まえた業績予測等
5	予実分析の自動化, アラート管理, PL 予測モデル作成
6	市場環境を踏まえた業績予測, 等
7	着地見込, 費用見込
8	将来予測のシミュレーション

出典：研究室の調査(2024)をもとに筆者作成

また、同質問票にて「予算プロセスにおいて AI をどのように利用または利用予定ですか」という問いを「予算管理のために AI・人工知能を搭載した会計システムを導入していますか」に対し、「導入予定」と回答した対象向けに質問した結果が図表 18 である。主に予測に関するタスクへの導入意向が見られ、市場環境等を踏まえた予測に利用することが目的である。

環境予測を支援することは人間よりも大量のデータを学習した上で予測を導き出せる AI の活用領域であるといえる。実態調査では将来環境の予測の関心が以前に比べて低下していることが明らかとなったが、今回の研究室の調査で AI を予算管理に導入するにあたって予測の支援が求められていることが判明した。

また実態調査からは榊谷(2022)の調査で以前と比較して予算数値と戦略とのリンクが高まってきていることが判明した。企業戦略・方針を AI に適切に反映させるアプローチとしては自然言語処理である LLM が適しているといえる。自然言語処理であれば仮に戦略を反映していないと思われる出力が返ってきたとしても、「今期の自社の戦略は〇〇であり、今出力された応答からはそれが考慮されていないように思えるので、もう一度演算

し直して」などといった具合にプロンプトから AI の制御が可能となるため、戦略の反映を適切に行えるのみならず、方針の変更時においてもプロンプトによって、モデルを組み直すことなく対処できるという強みを持つ。

従って、本論文の予算編成 AI は LLM によるアプローチを検討することとした。予算管理 LLM であれば AI を自動化ではなく、予算管理のためのアシスタントやコンサルティングといった要素で活用できるため、予算管理自動化上の逆機能も抑えることが可能となる。LLM では図表 16 予算管理システムについての課題として挙げられた「システム化」や「経理上の課題」を支援することはできないが、それ以外に寄せられた全ての課題に対してアプローチできる拡張性を持ったモデルを想定することができる。環境予測や市場予測について、自然言語処理のみでなくマルチモーダルモデルとして処理とすることにより、純粋な LLM と比較して精度向上を目指すことができる。予算管理 LLM は適切な処理が行えるよう開発工程の分割を行い処理層を分ける、「Base Model」と「予算管理 Model」の二層構造で AI を構築する。

5.1.1 BaseModel

Base Model とは翌期以降の財務諸表数値を予測することを目的とし、管理会計タスクを解く予算編成 AI の第一処理層として想定したモデルである。予測プロセスとしてテキストデータと数値データに時系列予測と LLM 処理を行うマルチモーダルシステムであり、財務諸表数値を合理的な水準で予測することで予算管理支援への第一歩とすることができる。

Base Model を活用することで、予算管理のみならず会計領域の LLM 活用における基礎的予測情報を導き出し、あらゆる会計予測 AI の基礎になりうる情報である財務諸表数値の予測を行う。

また、管理会計タスクを解く AI において翌期の会計数値予測に大きな乖離が生じた場合、AI の算出する回答の精度は低いものとなる。そのため、翌期以降の財務諸表数値を合理的な水準で予測する必要があり、モデルを分割して作成することにより会計数値の予測に特化した調整が行えるため、開発上の利便性に優れる。

Base Model で財務諸表の予測を実現するために、戦略・経済状況・市場状況・(政治)に関連する情報を学習(もしくはプロンプトで入力)させ、市場環境を予測した上で財務諸表

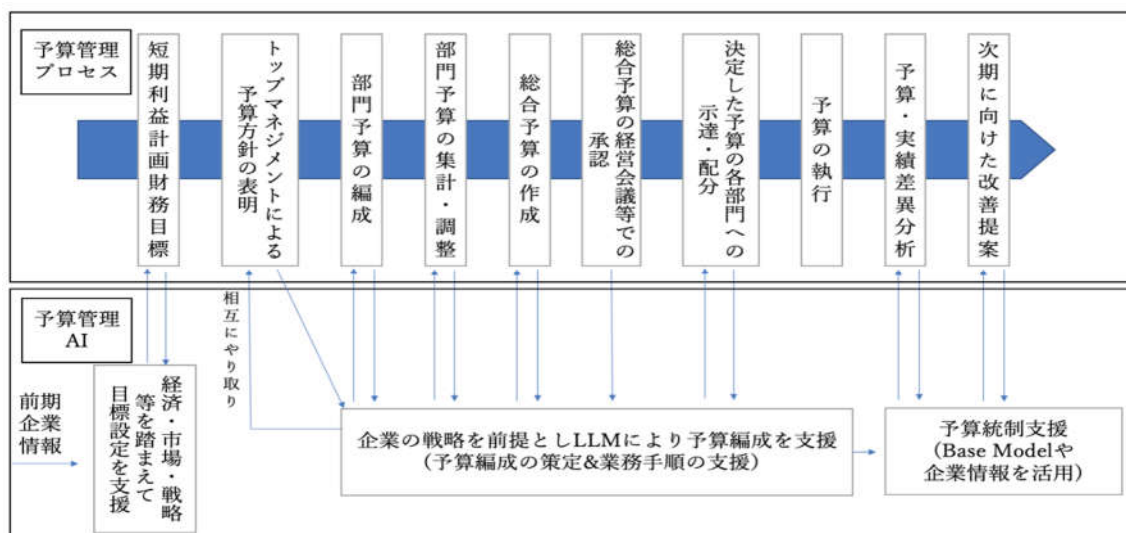
の予測に寄与させる。

5.1.2 予算管理 Model

ことで、予算管理の最適化を目指す AI システムである。戦略・経済状況・市場状況・政治といった要素は Base Model にてモデリングされているため、内部情報・他社動向・顧客情報・成長性に関する情報を追加的に学習させる。

図表 19 は予算管理プロセスに Base Model と予算管理 Model で構築された AI が支援を行う要素について図示したものである。

図表 19 予算管理プロセスと予算編成 AI の関係



出典:伊藤・日時(2021,p. 63)をもとに筆者作成

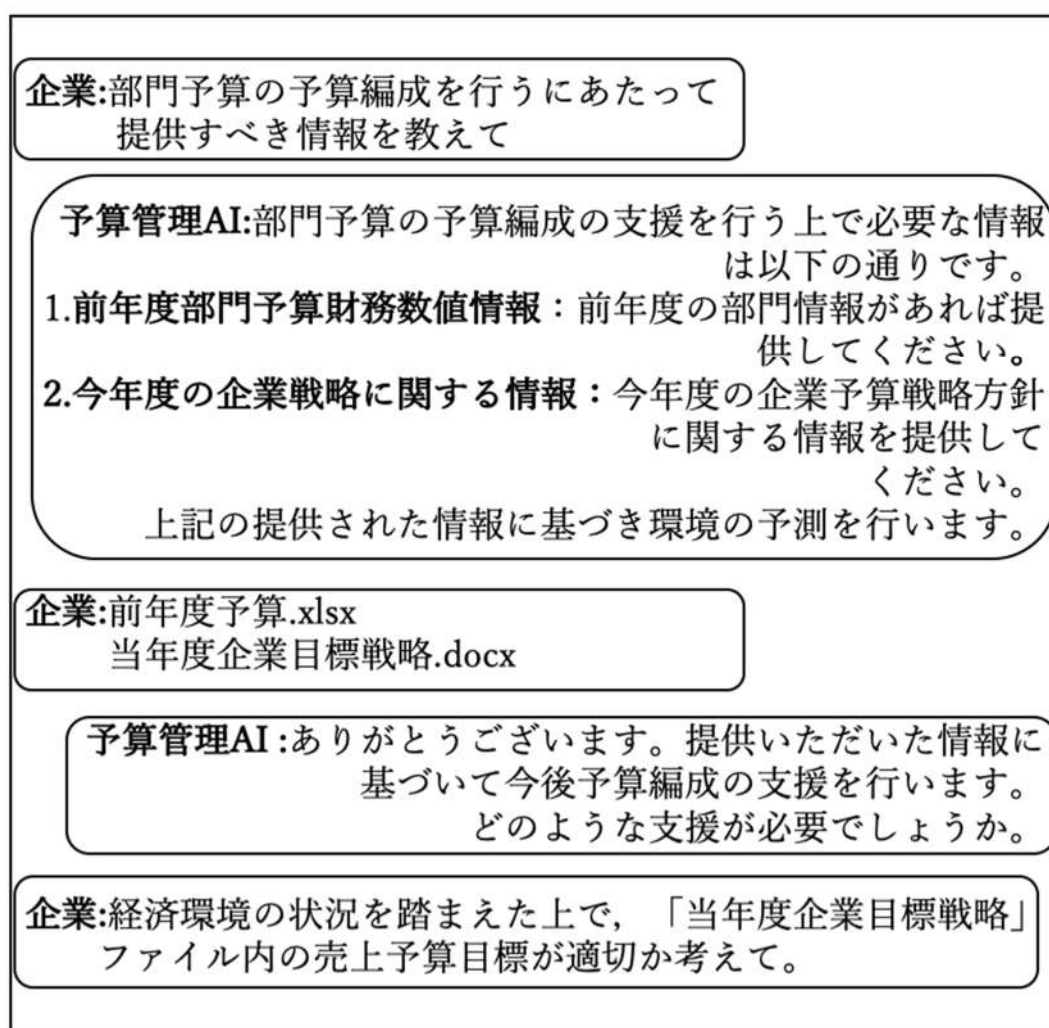
予算管理 Model は企業の戦略・内部情報・経済状況・市場状況・他社動向・政治・顧客情報・成長性といった管理会計に必要なデータを学習し企業の状況をモデリングする。

予算編成 AI は前期までの企業情報を学習しており、前期の予算統制に関わる事項も学習した状態である。短期利益計画目標の設定では企業の提示した目標に対し、前年度までの状況や現在の経済状況、当期の予測財務諸表数値から目標値に対してアドバイスを行う。トップマネジメントの方針の表明でも同じようにアドバイスを行い、場合によっては管理会計担当者からトップマネジメントへの提言を提案することで方針が誤った方向へ行くこ

とを防ぐ。予算の編成や調整といった段階でも支援を行い、時には予算管理行動の指示も行うことで適切な予算管理行動を行えるように企業を誘導する。予算編成が完了し、予算が策定された後も常に状況を把握し続けることで市場状況が変動した際にも迅速な対応がなせる。

予算統制プロセスでは差異が生じた要因をリアルタイム学習データから推測し、分析の支援を行う。期中の予算管理プロセスの情報や予算統制の情報を新たにデータとして学習することで、前年度よりも精度に優れた予算管理支援が行えることが推測される。

図表 20 予算管理 LLM イメージ図



出典：筆者作成

5.1.3 システムによる予算管理行動ナッジ

予算管理行動は一定規模以上の企業ではほぼ全ての企業が行なっていることが、先行研究から明らかとなったが、企業が適切な予算管理を行なっているかは不明である。適切な予算管理とは固定的な概念ではなく、企業ごとによる柔軟な対応を持って行うものだが、組織戦略が企業構造を規定し、それが会計システムの設計を順次形成されていくため (Dent 1990)、企業が予算管理を適切に行なっていなければ、その企業において設計された会計システムが企業に貢献することは困難だといえる。一般に企業戦略→会計システムの流れでシステムの構築は行われていくが、システム側から予算管理行動をナッジしていくことで、予算管理 LLM によって企業の予算管理行動をコントロールし適切な予算管理行動を取れるようすることが理想である。

システムファーストで企業内行動を構築した例としてミヤシステム株式会社の事例(菅本・宮脇 2020)では、日次レベルの損益管理システムを工程管理と実行予算管理を連動させたシステムを構築している。システムに求められる要件は、予算がプロジェクトの道標になり、設計書と現場を結びつけること、設計や積算の不備を早期に発見できること、工事担当者を動機付けできることが挙げられる。これらの要件を達成するために「作業のグループ化」による進捗管理や「対話型コントロール・システム」として現場状況の確認を毎日行うといった事例も存在しており、システムによって管理会計業務フローがナッジされた例だといえる。

予算編成 AI は予算管理プロセスの支援を行う LLM であるが、期待される効果は戦略を把握し、環境を予測することで予算編成のアドバイスを行うだけではなく、適切な予算管理プロセスを実施できるように支援を行う機能も有するべきである。具体的には管理会計担当者と予算編成 AI がプロンプトを通じてやりとりを行うわけであるが、AI システム側から企業側に対して情報提供を求めることによって、予算管理行動のナッジ機能の実現を図る。

5.2 予算編成 AI のための技術

第 3 章にて本論文における AI の定義を行うために AI の変遷の解説を行ったが、予算編成 AI に使われるアルゴリズム理解のために必要な技術があるため、予算編成 AI に当該技術がどのように貢献するかといった観点から AI の技術の検討を行う。本節で説明を行う

技術によって予算編成 AI は予算管理の予算編成をはじめとした機能を支援し、企業の予算管理行動を適切にナッジする予算管理支援 LLM として機能することができる。

5.2.1 予算編成 AI の教師データ

AI システムの構築のためには AI に学習させるための教師データと呼ばれるデータを学習させる必要がある。AI においてデータの確保は大きな障壁となる。データとは、さまざまなソースから得られるものである。過去の財務データの時系列分析など、一般的にクリーンで信頼できると考えられている組織自体の記録からもたらされることが主流であるといえるが、商業、ソーシャルメディア、業務、公共、他企業のデータなど、外部へのアプローチでの入手も可能である。その他にも「生で洗練されていない」状態のデータとして、ユーザーのクラウドデータ、インターネット上のその他の媒体から収集されたデータ、ノイズとなるデータ、未精製のデータが存在する。(Moffitt and Vasarhelyi 2013)しかし、企業にとって「生で洗練されていない」状態のデータは有用ではないため、意味のある答えを導き出すために分析する必要がある。(Al-Htaybat and von Alberti-Alhtaybat 2017)

データには量と質が求められ、管理会計のような内部的な領域においては特にその問題が大きいといえる。筆者の知る限りでは汎用的管理会計タスクに対応する AI の教師データを入手可能性も踏まえた上で検討した研究は存在しないため、管理会計の中でも予算管理タスクに対する教師データの検討を行う。教師データには量と質が求められるため一定のルールが設定されているデータが望ましい。よって教師データとしては「有価証券報告書」「四半期報告書」「決算短信」「統合報告書」などの開示情報や「アナリストレポート」「ニュースデータ」「SNS」といった第三者からの情報、「企業の予算編成データ」「企業の予算統制データ」「関連議事録データ」などの企業内部データが候補となる。

有価証券報告書

有価証券報告書は特定のフォーマットによる開示がされており、監査による質の担保も行われているため情報の量、質ともに優れたデータであるといえる。資料内には企業の基本情報や事業内容などの「企業概要情報」、事業方針やリスク管理、市場環境への対応などの「経営者の分析情報」、損益計算書や貸借対照表などの「財務諸表情報」が含まれている。「企業概要情報」や「経営者の分析情報」はテキスト情報「財務諸表情報」は数値情報が

主体となるため、AI モデル構築上の扱いが異なることとなる。いずれのデータも適切に前処理を施すことで予測に対し有用であることが推測される。

決算短信

決算短信には売上高や利益の経営者予測情報が含まれており、全社的予算管理のデータに近いものであると推測できるため、教師データとして活用する。経営者予測情報は困窮している企業が good news を出した場合、そうでない企業に比べ誤差が大きいことや経営者は売上に対し楽観的・強気な傾向にあること、業種によって予測の精度に差があることが明らかとなっているため、活用の際にはその特徴を考慮することが重要である。(西・金田 2009; 広原 2021) 決算短信による経営者予測情報は企業内部データである管理会計情報を利用できる際には不要であると考えられるが、企業の内部データは入手が困難であるため、代替的データとして用いている。

ニュースデータ

経済ニュースやビジネスニュースを教師データとすることで経済の流れや企業の業績、企業の新製品、政治動向を LLM に学習させることが可能となる。この学習させた知識をもとにマルチモーダルを活用により財務諸表の数値データを補完することが可能である。そのようにモデルを構築することで財務諸表数値の深い解釈が実現し、AI モデルの予測精度向上につながると考える。財務諸表数値の補完情報としては企業の内部データを用いるのが最も効果的であると思われるが、入手可能性の問題があることや内部データのみを用いると客観性に欠くことから第三者的情報であるニュースデータを利用する価値は高い。

企業会計基準、財務諸表規則

事前学習済み LLM では会計に関する知識はある程度学習されているが会計タスクに特化させるには十分とはいえない。GPT-4 を持ってしても執筆時点では日本の公認会計士試験を解くほどの知識や会計に対する思考力はなく仕訳問題はほとんど解くことができなかった。また類似した知識である IFRS や US-GAAP 等の日本以外の会計基準も学習していることから知識の混在の可能性も秘めている。有価証券報告書や会計関連ニュースなどを正確に学習させるためにも現行の企業会計基準や財務諸表規則を学習させることが有効に働く。

上記4つの教師データは企業が有する内部情報と比較して入手可能性に優れており、かつ一定の質の確保もなされるため、予算編成 AI 開発の上で適したデータセットとなる。精度確保ためには欠かせないデータの入手であるが、AI に教師データを作成させそれを学習させるというアプローチも存在する。Nishi et al.(2021)ではニュース記事を用いて株価予測を行うモデルのニュース取得制限問題を解決する策として GPT-2 を使用してニュースを作成し、それを予測モデルの教師データとすることで精度を向上させた。管理会計 AI にしても企業の内部データは有限であることから、AI に仮定の企業内部データ作成させ学習させることは有用なアプローチである可能性がある。

5.2.2 ファインチューニング

ファインチューニングとは、事前に大量のデータで学習済みの LLM に対して、特定のタスクやデータセットに合わせた追加の学習を行う手法である (Church 2021)。このプロセスは、元のモデルが持つ一般的な知識を特定の用途に特化させることを目的としている。よってこの手法を用いて学習済みの LLM に対して管理会計タスクをこなすための知識を学習させることとなる。この手法によって元のパラメーターを保持したまま新しいデータに適用できるため、従来のモデルの知識を活用することができることや、学習ゼロからモデルを構築する際に比べ学習時間を短縮できるといったメリットもある。

5.2.3 RLHF

人間の好みに合わせた言語モデル (LM) の調整手法として、RLHF (Reinforcement Learning from Human Feedback) が注目されている。RLHF は、強化学習を基礎としており、人間からのフィードバックに基づいて報酬関数を学習する (Christiano et al. 2023)。強化学習は、試行錯誤を通じて価値を最大化する学習手法であり、これを用いることで、たとえ事後学習モデルが小さくても LLM からの出力を優先することが可能となる (Richard and Andrew 2019; Ouyang et al. 2022)。RLHF を LLM に適用することで、システム設計者の望む返答に近づけることが期待される。本論文では管理会計に特化する LLM の構築を目的の一つとしているが、パラメーター数に優れたモデルは様々な情報の教師データを学習しているため、利用する LLM には会計情報以外の教師データが含まれている。中には金融に特化したモデルなども存在するが (Wu et al. 2023)、言語モデルは学習した情報から

学習していない情報の回答を導き出す能力(Commonsense)を有していることが GPT-2(Radford et al.2019)の研究により判明しているため、会計特化 LLM であるとしても様々な情報を学習している LLM を使用するのが望ましいと考えられる。これは機械学習や統計学で用いられてきたオッカムの剃刀のような考え方とは異なるものとなる。RLHF は様々な情報を有する LLM に管理会計情報を齟齬なく回答させる手段の 1 つとなり、アプローチを検討する。

RLHF を実施するためには LLM 構築後の回答に対し、手動でフィードバックを行う必要がある。管理会計に関連する回答を選択することでモデルをタスクに特化させることが可能となる。行う必要のあるフィードバック量はモデルのパラメーター数や行いたいタスクにより変動する。次にフィードバックに基づく報酬関数の定義を行う。人間のフィードバックに基づく報酬関数の定義の例は以下のとおりである。

$$P(A \succ B) = \frac{e^{r(A)}}{e^{r(A)} + e^{r(B)}} \quad (1)$$

式 1 は B の回答よりも A の回答が優先される場合の式であり、 $r(A)$ と $r(B)$ はそれぞれの回答に対する報酬関数の推定値である。A と B の回答の好みを相対的な確率で表すことにより、報酬信号(どれだけその回答が優れているかを表す数値的指標)として利用しモデルの出力を方策勾配法でパラメーターの重みを調整する θ を算出する。

$$\theta \leftarrow \theta + \alpha \nabla_{\theta} \log \pi_{\theta}(a|s) G_t \quad (2)$$

θ = LLM の重み

α = 式 (1) ($A \succ B$) で算出した報酬信号(学習率)

s = 状態, モデルが決定を下す際の状況や文脈を表す。具体的にはユーザーからの入力を固有な分散表現に変換したものとなる。

$\log \pi_{\theta}(a|s)$ = 確率のパラメーター θ による勾配。(対数変換しているのは計算簡略化のため)

G_t = モデルの回答

a に対する報酬の合計

式 1 によって LLM に対して回答 A の評価や回答 B の評価がフィードバックされ学習される。式 2 は基本的な方策勾配法を LLM に対応させたものであり人間のフィードバックを LLM に学習させることは可能であるが、方策の更新が不安定であるという問題点を抱えており、近年用いられるモデルでは方策の更新に工夫を凝らしたモデルが用いられていることで問題点が改善されている。

管理会計に特化した LLM では管理会計に関する回答のみを期待することから、RLHF を用いたファインチューニングを通じて管理会計タスクに特化した LLM の構築につなげることができる。

5.2.4 プロンプトエンジニアリング

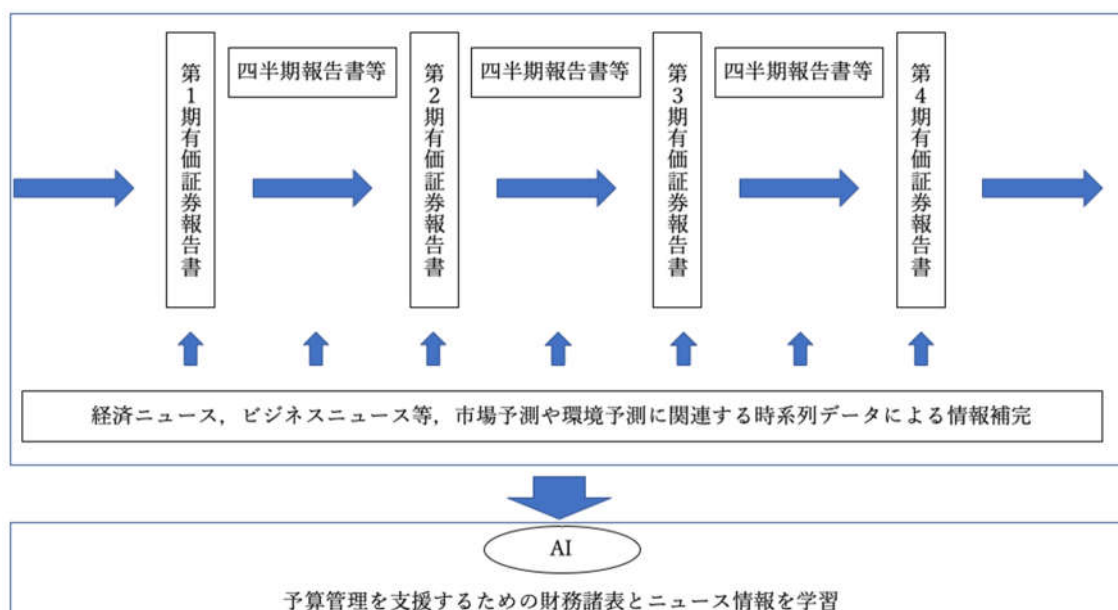
LLM にタスクを実行させるための入力をプロンプトといい、プロンプトの設計により LLM の解答精度に差が生じる。管理会計 AI は企業の予算管理プロセスをアシストする目的で構築されているため、精度の高い回答を追求する必要がある。例えば、プロンプトに回答例を提示する Few-Shot(Brown et al.2020)を用いることによって、回答例がない状態 (Zero-Shot) よりも回答精度が向上することが判明している。Few-shot の精度はモデルが大規模なほど精度が向上するケースが多く、文脈内から学習するため In-Context Learning と呼ばれる。Few-shot の際に思考の過程を入れる Chain-of-Thought(Wei et al.2022)では推論能力を向上させることができ、管理会計のタスクを解く上でも有意な手法である。プロンプトエンジニアリングから判明した LLM の変わった挙動としては LLM を応援するプロンプトを入力するとタスクの解決力が向上といったこと (Li et al.2023) や LLM に報酬を与える旨をプロンプトに加えると実際に報酬を与えなくともタスク解決力が向上する (Lightman et al.2023) といった事例もある。LLM のアルゴリズムはブラックボックス化しており、Transformer 系統のモデルである以上は挙動の説明が容易ではないため、予算編成 AI のためのプロンプトフォーマットを画一化することは困難だといえる。様々な既出の手法に倣ってプロンプト実施をすることは管理会計タスクの解決に繋げる一つの手段であることは間違いないと考えられるが、実務において一定の精度を確立するためにはシステム利用者の情報収集による検証が必要となる。

5.2.5 時系列データと予算編成 AI

予算管理システムの構築において、LLM(Large Language Models)を活用する場合、時系列情報を含むニュースなどのデータを予測モデルに反映させることにより、モデルの精度向上が期待される。有価証券報告書は、四半期ごとに提出されるものを含め、通常 3 ヶ月に 1 度の頻度で提出される。この報告書に記載されている数値の変動理由を予測する上で、ニュースデータなどを利用することが有効であると考えられる。例えば、1 月に提出され

た四半期報告書と4月に提出された四半期報告書の間が発生した数値の変動を分析する際、1月から4月までの間に発表されたニュースを基に経済動向を把握し、それを予測に反映させることで、精度の向上が見込まれる。このような分析を効果的に行うためには、LLMに時系列の特性を組み込むことが重要であり、そのようなアプローチにより、予測の精度向上が期待される。

図表 21 有価証券報告書情報を時系列情報で補完するイメージ図



出典：筆者作成

NLP タスクでは1つのモデルで複数のタスクに対応できるが、時系列タスクは多くの場合特化されており、タスクやアプリケーションによって異なるデザインが必要である。例えば時系列予測は、需要計画(Seyedan and Mafakheri 2020)や在庫最適化(Namir et al. 2022)や気候モデリング(Amato et al. 2020)に至るまで、多くの実世界の動的システムにおいて重要な能力である。それぞれの時系列予測タスクは、通常、広範な専門知識とタスク固有のモデル設計を必要とする(Jin et al.2023)。

5.2.6 LLMTime

予算編成 AI の Base Model では Gruver et.al(2023)が Github で公開している LLMTime

というモデルを利用している。LLMTime は時系列データを数値の羅列として符号化することで、時系列予測をテキストにおける次トークンの予測として捉える技術を発展させることにより、LLM が特定のタスクで訓練された専用の時系列モデルに匹敵する精度をだすアプローチである。時系列予測を次トークンの予測と見做すことにより、事前訓練された LLM と尤度評価やサンプリングなどの確率能力の利用を可能とする。このモデルの利用によって数値データとテキストデータが混在する有価証券報告書から粒度を揃えた予測を実現可能とする。LLM と時系列予測を比較すると両者は類似性があり、同様のシーケンスの集合であることがわかるが、LLM はテキストデータを高次元ベクトル空間への埋め込みによって処理し、これらのベクトル表現は単語やフレーズの複雑な意味的關係を捉える。この際、各ベクトル成分の数値自体には直接的な意味はなく、それらは合わせてテキストの意味合いを符号化する。これに対し、時系列予測モデルでは実際の数値データ(例えば経済指標、気象データなど)を直接処理し、これらの数値は実測値や観測値としての明確な意味を持つ。したがって、これら二つのモデルの出力は、その性質と解釈方法において顕著な違いがある。さらに LLM では単語や文字である自然言語を扱うことからデータ分布は原則として離散的なものとなる。一方時系列データは時間とともに変化するデータを扱うため、連続的な数値として表されるケースも存在する。LLMTime ではトークン上の離散分布を連続密度関数へ変換することができ、時系列予測タスクにおいて従来では困難だった分布やモダリティを扱うことが可能となる。

---LLMTime を予算編成 AI に適用することによる利点---

①事前学習された LLM と財務諸表の時系列情報を並列に処理し予測に寄与することが可能である。

LLMTime を利用することによって時系列予測を自然言語処理に近い形式で処理することが可能となるため、実質的に時系列予測と自然言語処理のマルチモーダル化が実現する。仮に LLMTime を利用せずに時系列予測情報を LLM に組み込もうとする場合プロセスが多岐に渡ることとなる。時系列予測用の前処理を行い、ARIMA(Auto-Regressive Integrated Moving Average), LSTM (Long Short-Term Memory), GRU(Gated Recurrent Unit), RNN(Recurrent Neural Networks)などの時系列用モデルを選択しハイパーパラメーターチューニングを行う。その後 LLM に学習させるのに最適な形式に前処理を行いテキストデータと共に LLM にファインチューニングさせることとなる。LLMTime を利用した場

合は時系列データとテキストデータを同時に学習させるため、ハイパーパラメーターの調整が非常にやりやすくなる。さらに時系列データの予測後データを別途前処理する必要がなく、モデルによって時系列情報の処理が最適化されていることから簡易かつ高精度な予測が実現可能となる。

②少量の時系列データであっても LLM の情報による補完が可能である。

LLMTime は LLM を直接的に学習するモデルではなく、既存の LLM を利用し時系列データを処理するための手法である。時系列データとして財務諸表数値データを利用するが期末時点や四半期ごとの数値情報では一般的な時系列モデルにおいて圧倒的にデータが不足しているといえる。当該モデルのメリットは5年分といった短い期間のデータでも LLM のテキスト情報を利用することで有用な予測を行うことができる可能性がある。LLMTime は LLM と結合し予測を行うモデルであるため、財務諸表数値データの説明要因となる有価証券報告書のテキストデータや経済事象や動向を踏まえるためのテキストデータを結合先 LLM に追加的な学習をさせる必要がある。

③LLM はインピュテーションなしで欠損値を扱うことが出来るため財務諸表データに強い。

本来、統計モデルや機械学習モデルでは欠損値がある状態で予測や分析を行うことはせず尤もらしい数値で補完を行うか、欠損が生じているデータは削除する前処理を行う。しかし、今回扱うデータは財務諸表データであるため、企業ごとに扱う勘定科目の名称は異なる点や特定の業種のみ存在する勘定科目や個々の企業特有の勘定科目もあり欠損値が存在することが当然である。このような性質を持つデータは従来の時系列モデルでは扱うことが困難であったが LLM では欠損を特定のトークンに置き換えることによって欠損自体が意味をもつシグナルとして機能し、財務諸表データのようなデータでも適切な予測に繋げることが可能となる。

LLMTime はメリットだけでなく予算編成 AI との相性がよくない要素も存在する。Base Model は次期以降の財務諸表項目の数値を合理的な水準で予測することを目的としたモデルであり、予測の精度を確保するには追加的に学習させる関連情報を強く考慮させる必要がある。RLHF やファインチューニングによって学習させることは可能であるが、そういった人の手の介在による手法は LLMTime に悪影響与えることが判明しており、トレード

オフの関係となる。

5.2.7 Meta Transformer

今回の予算編成 AI では適用しなかったが、管理会計の AI 適用時に有用と考えられるモデルの検討を行う。十数種類のモダリティからのデータを同じパラメータセットを用いて同時にエンコードするマルチモーダルフレームワーク「Meta Transformer」を紹介する。(Zhang et al.2023) Meta Transformer は Frozen Encoder を利用してペアとなるマルチモーダルを学習なしで実行する。Frozen Encoder とは前述した Encoder に工夫を凝らしたものであり、Encoder の重みを固定しトレーニングの間更新しない状態を指す。一般に Frozen Encoder を使うことで大規模モデルのパラメータトレーニング時に必要な部分のみを更新し、効率的なトレーニングが可能となることや Encoder の固定化によって学習プロセスが安定し過学習リスクを減らすために用いられるが、Meta Transformer では異なるモダリティ(画像、テキスト、音声など)の中で共通の特徴表現を獲得することに重点を置いている。共通の特徴抽出を可能とするのは事前に学習したデータからエンコーダーが一般的な特徴を有しているためであり、異なるタイプのデータに対しても共通トークン空間へのマッピングを通じて、基本的な特徴抽出メカニズムによって統一された理解と処理が可能となる。

Meta Transformer で扱えるデータの種類は「自然言語」「画像」「音声」「赤外線」「動画」「テーブルデータ」「時系列」「分子グラフ」「ハイパースペクトル」「慣性計測装置(IMU)」「X線」「3D vision」であり、非常に多岐なデータを同時に扱うことができる。仮に教師データに制約がないとすれば予算管理のみならず、管理会計のあらゆる論点に対応できるモデルを構成できる可能性を秘めている。LLMTime では上記データのうち、自然言語と時系列のみを扱った手法となるが、画像が扱えれば統合報告書そのものの利用や製品の学習などが可能となり、動画情報が扱えれば感情分析を用いた社内の風通しの活用や BtoC であれば顧客の映像情報を学習させられる可能性がある。音声データでは社内会議の直接的な学習が可能であり、テーブルデータは社内で作成したテーブルデータの活用や excel データを活用することができる。専門的な業種であれば「分子グラフ」「ハイパースペクトル」「慣性計測装置(IMU)」「X線」も利用できるだろう。これらの要素からの学習が実現すれば経済実態や企業状況、戦略の適切な反映に繋がり、意思決定会計や戦略策定原価企画な

どの領域でもアシスタントが可能となるモデルの構築も考えられる。

第6章 予算編成 AI の開発の流れ

第5章で予算編成 AI の構想をおこなったが、実際の開発において複数の問題点や今後普及していくにあたり障害となる事項が明らかとなった。現在、黎明期にある管理会計の AI 開発において障害となった事項を明記することで、将来の管理会計 AI 開発における指標となることが本研究の貢献の1つとなる。

AI を開発するにあたって経るプロセスは次の通りである 1. 「設計の考案」 2. 「データの収集」 3. 「前処理」 4. 「学習」 5. 「チューニング等の調整」 6. 「完成(以後も調整が必要)」。

なお、本論文においてシステムの完成には至らなかったため、精度の検証の実施は叶わず調整や完成の段階の障壁については把握できていない点は今後の課題とするところである。

6.1 設計の考案

このプロセスでは AI プロジェクトの要件、目的を明らかとし問題解決へと導くための指針策定プロセスである。本論文では先行研究の実態調査の結果から予算管理プロセスに求められている要件を把握し、予算管理プロセスの長所を排除しないことを意識した構想を行った。予算管理プロセスで生じている問題に当てはまる技術を AI の先行研究から探し適合する技術を利用した結果 LLMTime を主軸として構成することとした。仮に人的、金銭的資源が確保できる場合はトークナイザーやアルゴリズムの開発を検討することで精度向上に繋がることが考えられる。

要件定義の方針が誤っている場合その後の工程が全て無駄になってしまうため、十分に検討を重ねて、開発方針の決定を行うべきである。

6.2 データの収集

利用する教師データは 5.2.1 「予算編成 AI の教師データ」で検討した通りであるが、比較的入手可能性の高いデータであっても AI の学習のためには大量のデータを必要とする

ため、手間や権利上の問題が生じる。

図表 22 予算管理 AI システムの導入状況

Q14. 予算管理のために AI・人工知能を搭載した会計システムを導入していますか。	
導入予定	10 社(8.3%)
導入予定なし	110 社(91.7%)
導入済み	0 社(0%)

出典：研究室の調査(2024)をもとに筆者作成

今回研究室で行なった調査で予算管理のために AI・人工知能を搭載した会計システムを導入しているかを尋ねたところ、予算管理に AI システムを利用している企業は 1 社もないという結果となった。この結果は事業会社において AI を活用できるだけの管理会計データを企業が有していないことや、予算管理に AI 投資をかけるほどの関心がないことを表しており、今後予算管理や管理会計の AI の普及の上で障壁となる可能性が高い。導入予定の企業が非常に少ないことから AI システムを活用するインフラが整っていない可能性を示唆しており、多種の企業の管理会計データを学習し管理会計の正解を導き出そうとする管理会計 AI では取り組む企業が少なれば精度の確保も難しいため、現状のままでは教師あり学習を主軸とした AI 手法を定着させることは困難だといえる。

有価証券報告書

有価証券報告書のデータは金融庁の公開する EDINET を通じて取得することで大量の有価証券報告書を大量に取得することが可能となる。EDINET が公開する書類一覧 API を利用することで EDINET から有価証券報告書の PDF データや XBRL の取得を実現する。XBRL は事業報告の情報を分析・比較するために提供されている構造化された情報であるが、構造化されたデータは数値データがメインであり、テキストデータは企業の記載した情報がそのまま抽出されるため、自然言語処理では扱いづらく LLM へのファインチューニングデータとするための前処理が困難であった。加えて XBRL はタクソノミによって財務諸表の要素をタグづけしたデータであるが、XBRL のままでは AI に学習させることができなため、csv, xml, json, nosql などの形式に変換し学習させる必要があるため付与されているメタデータを利用するためには工夫が必要である。また、本論文における有価

証券報告書は日本基準適用会社のみを対象とした。IFRS や USGAAP の下で作成された財務諸表は数値の持つ意味が変わってしまい学習時に齟齬が生じるためである。

決算短信

決算短信データは執筆時点では EDINET の書類一覧 API に該当しないため、TDnetAPI を利用することで XBRL 形式での情報取得が可能となる。もしくは EDINET や TDnet から PDF を取得する方法でも経営者予測情報を抽出することで利用可能である。あくまで一例ではあるが、PDF パーサー(python の pyPDF2 や PDFMiner など)を利用しテキスト情報の抽出をし、定型的なフォーマットであるためテキストから情報抽出ルールを定義することによって経営者予測情報の利用が可能となる。

ニュースデータ

従来ニュースデータを教師データとする場合、センチメント分析や LDA, word2vec のような手法で文の一部の特徴を捉えた学習を行っていたが、Transformer モデルの場合ニュース本文の全体を対象として学習させることができ、文脈の考慮が可能となることからより高度な学習が可能となる。

ニュースデータの大量入手の方法としてはニュースサイトの提供している API の利用、スクレイピングの利用、ニュースサイトとの契約を通じて提供などの手法があるが、昨今の生成 AI ブームを通じて多くのニュースサイトはニュースの商用、非商用を問わず AI 利用を禁止している。ニューヨークタイムズが Open AI に対しニュース情報を利用したとして訴訟を起こしているように(日本経済新聞 2023 年 12 月 23 日, 日経速報ニュースアーカイブ), ニュース提供サイトの規約を理解した上で学習させる必要がある。現在の AI の学習に対する排他的姿勢を鑑みるとニュースサイトとの契約を行いニュース情報の提供をしていただくという手段が主流となるだろう。しかし, AI に学習させるデータ量は非常に多く, 金銭的問題が開発の上で障壁となる可能性がある。前述した通り本論文では TimeLLM の利用により, 財務諸表の数値変動をニュースデータで意味補完し予測に役立てることを目的としてニュースデータを利用しているが, 経済状況の把握が重要である管理会計の AI 構築において, ニュースデータをリアルタイムで AI に学習させる状況を生み出すことは非常に重要であると考えられる。

企業会計基準，財務諸表規則

ASBJ や金融庁から PDF を取得し学習。

その他のデータ

その他の検討事項のデータとして「四半期報告書」「統合報告書」「アナリストレポート」「SNS」「企業の予算編成データ」「企業の予算統制データ」「関連議事録データ」などが挙げられる。「四半期報告書」は純粹に有価証券報告書のデータを増加させる効果があるが、本論文では処理にかかる時間を考慮して除外することとした。「統合報告書」は企業が自由なフォーマットで制作しており、情報の抽出が困難である。試しに筆者が日清の統合報告書を GPT4V で読み取り説明をさせたところ満足いく結果とはならなかったため、現時点における活用は難しい。「アナリストレポート」は投資 AI など教師データとされることがあるが(鈴木ほか 2022)内部情報を扱う管理会計の領域においてどれほどの有用性があるのかは未知数である。「SNS」は予算編成への AI アプローチを検討した Marotta and Au (2021)でも取り上げられており、SNS 情報の利用により個々のステークホルダーの重要性和チャンネルの把握に繋がるとした。しかし、他の教師データと比較した際に質の低いデータであり、信憑性に欠けることから利用には精査が必要となるだろう。

上記データ群は相対的に入手性に優れており、予算編成 AI の「Base Model」で主に利用できるデータであるが「予算編成 AI」の主要となるデータは特に入手が困難である。「企業の予算編成データ」「企業の予算統制データ」「関連議事録データ」は予算管理を支援するために必須のデータであり、多種多様な企業からデータを収集し学習させることが望ましいがその手段は困難を極める。さらに、有価証券報告書などの外部報告向け資料と異なり、内部管理用のデータはフォーマットが定まっておらず企業によって力の入れ具合も異なってくるため、前処理もルールベース化できない箇所があり人の手で行う必要が出てくる可能性がある。その場合は管理会計を理解した上でデータ処理ができる人材を集めなくてはならない。よって、企業内部データ利用の際のイニシャルコストは非常に高く費用対効果に見合うかが内部情報を利用する管理会計 AI の課題点となるだろう。

6.3 前処理

図表 23 は有価証券報告書を EDINET から取得し、全ての有価証券報告書の項目を csv

化した表の一部である。数値データとテキストデータが共存していることや欠損値が存在していることが把握できるため、データを区別して処理を行うことが望ましい。テキストデータと財務諸表数値を複合して利用することは、単にどちらかのみを利用したモデリングよりも精度が向上することが財務会計の領域では実証されている。Wu and Du(2022)は LSTM や GRU を利用して MD&A から中国上場企業の財務諸表不正の検知に関する実証研究をおこなった。結果としてテキストデータと財務諸表数値両方を利用したモデルの精度が一番優れており、テキストデータのみを利用したモデルの一番精度が低かった。

図表 23 XBRL 有価証券報告書データ

売掛金	貸倒引当金、流動資産、一括控除	貸倒引当金、投資その他の資産、一括控除	主要な販売費及び一般管理費 (テキストブロック)	一般管理費
51000000	-49000000	-24000000	※2、一般管理費のうち主要な費目及び金額は、次のとおりであります。前事業年度(自平成28年4月1日)に	
8782000000	0	-24000000	※2、販売費及び一般管理費のうち主要な費目及び金額は次のとおりであります。前連結会計年度(自平成27年4月1日)に	
2.09683E+11	-811000000	-1229000000	※2 販売費に属する費用のおよその割合は前事業年度60%、当事業年度60%、一般管理費に属(①)※1 一	
2840000000	-453000000	-92000000	※1 販売費及び一般管理費のうち主要な費目及び金額は次のとおりであります。前連結会計年度(自平成27年4月1日)に	
	-20000000	-10000000		
	-22560000	-129000000	※2 販売費及び一般管理費のうち主要な費目及び金額並びにおよその割合は、次のとおりであります。前連結会計年度(自平成27年4月1日)に	
6107000000		-76000000	※1 販売費及び一般管理費のうち主要な費目及び金額は次のとおりであります。当連結会計年度(自平成28年4月1日)に	
97959000	-94908000	-29991000		
799557000	-317000	-125975000	※2 販売費及び一般管理費のうち主要な費目及び金額並びにおよその割合は、次のとおりであります。前連結会計年度(自平成27年4月1日)に	
1362079000	-32505000	-9485000	※2 販売費及び一般管理費のうち主要な費目及び金額並びにおよその割合は、次のとおりであります。前連結会計年度(自平成27年4月1日)に	
	-39000000	-218000000	※2 販売費及び一般管理費のうち主要な費目と金額は次のとおりであります。前連結会計年度(自平成27年4月1日)に	
			※2、一般管理費に属する費用の割合は前事業年度100%、当事業年度100%であります。営業費	
9762400000	-204000	-3149000	※2 販売費及び一般管理費のうち主要なもの 前連結会計年度(自平成28年4月1日)至平成29(※3)研究	
17019000000	-27000000	-298000000		
137660000				※4、一
	-274000	-136114000		
426000000	-1776000000	-15000000	※1 販売費及び一般管理費のうち主要な費目及び金額は次のとおりであります。前連結会計年度(自平成27年4月1日)に	
	-6391000	-27873000		
10311000000	-44000000	0	※1、販売費及び一般管理費のうち主要な費目及び金額は次のとおりであります。前連結会計年度(自平成27年4月1日)に	
3302539000	-1966000		※2、販売費に属する費用のおよその割合は前事業年度55%、当事業年度55%、一般管理費に属(※3)研究	
	-226000000	-116000000	※2 販売費及び一般管理費のうち販売費に属する費用のおよその割合は前事業年度69%、一般	

出典：筆者作成

数値データ

数値データの処理として機械学習の処理や統計の処理などでは欠損数値の補完を行うインピュテーションアプローチがよく取られるが、それは欠損値が計測上の問題からデータの取得ができなかったために生じるものであり、その補完として行われるデータの前処理手法である。有価証券報告書において欠損値はデータ集計上の問題ではなく、企業の記載事項ではない項目であることからインピュテーションを実施できない。LLM では欠損を特殊なトークンに置き換えることで、欠損のまま予測の実現が可能となる Gruver et al.(2023)。

テキストデータ

Transformer モデルの自然言語処理における前処理は基本的に以下のプロセスを経ることとなる。

①データクレンジング

テキストデータから不要な文字を削除する。

②データの正規化

テキスト形式の統一

③テキストのトークン化

テキストを数値表現にするためにテキストを小さな単位へと分割する。

④特殊トークンの追加

Transformer モデル用の特殊トークンを追加する。

⑤単語分散表現

トークン化された単語を埋め込み層によってベクトル(テンソル)表現に変換する。

⑥パディング, トランケーション, アテンションマスク

テンソル表現された単語のシーケンスを調整する。

しかし, XBRL から抽出されるテキストデータは自然言語処理での利用は想定されておらず, 非常に前処理が難しいデータであった。図表 24 は有価証券報告書の連結経営指標等の一部である。

図表 24 連結経営指標等

連結経営指標等						
(1) 連結経営指標等 回次	第15期	第16期	第17期	第18期	第19期	決算年月
平成26年3月	平成27年3月	平成28年3月	平成29年3月	平成30年3月	売上高 (千円)	— — 12,086,493 13,961,972 17,033,374
— —	822,054	804,406	574,445	— —	経常利益 (千円)	— — 553,498 548,871 349,469
— —	553,498	548,871	349,469	— —	親会社株主に帰属する当期純利益 (千円)	— — 553,498 548,690 351,316
— —	4,424,059	7,609,200	7,889,655	— —	純資産額 (千円)	— — 18,588,068 26,005,361 26,111,454
— —	127.46	202.26	209.03	— —	1株当たり純資産額 (円)	— — 15.95 15.74 9.29
— —	23.8	29.3	30.1	— —	1株当たり当期純利益 (円)	— — 13.2 9.1 4.5
— —	86.6	58.2	81.7	— —	自己資本比率 (%)	— — 13.2 9.1 4.5
— —	1,547,242	1,852,721	3,067,888	— —	営業活動によるキャッシュ・フロー (千円)	— — 1,547,242 1,852,721 3,067,888
— —	△2,550,062	△5,410,917	△1,382,827	— —	投資活動によるキャッシュ・フロー (千円)	— — 1,940,835 4,349,483 △1,941,020
— —	4,077,672	4,868,773	4,612,747	— —	財務活動によるキャッシュ・フロー (千円)	— — 4,077,672 4,868,773 4,612,747
— —	339	495	563	— —	現金及び現金同等物の期末残高 (千円)	— — 339 495 563
— —	—	—	—	— —	従業員数 (名)	— — [ほか、平均臨時雇用人員] [—] [—] [51] [79] [72]
— —	—	—	—	— —	(注) 1	第17期より連結財務諸表を作成しているため、第16期以前については記載しておりません。
— —	—	—	—	— —	2	売上高には、消費税等は含まれておりません。
— —	—	—	—	— —	3	平成27年9月1日付けで普通株式1株につき普通株式4株の割合で株式分割を行っております。第17期の期首に当該株式分割が行われたと仮定し、1株当たり純資産額及び1株当たり当期純利益を算定しております。
— —	—	—	—	— —	4	潜在株式調整後1株当たり当期純利益については、潜在株式が存在しないため記載しておりません。

出典：⁶筆者作成

図表 24 は 1 つのコンマ区切りの中に決算年度に準じた数値が連続して並んでおり、それぞれの年度との対応関係を反映させることが難しい。さらに単位が(千円)(%)のように並んでおり、自然言語と数値が混在しているため数値をコンピューターに読み込ませる上では適切に変換する必要がある。また注釈の情報を学習時にいかに反映するかというのも課題となる。

このようなテキスト情報が多くの項目で存在し、全ての日本基準適用会社である上場企

⁶ 図表の連結経営指標等はさくらインターネット株式会社第19期有価証券報告書のXBRLをもとに作成

業を対象としたデータ集計を行なっているため、手作業による前処理は実質的に不可能である。有価証券報告書で提出される情報であるため、ある程度画一的なフォーマットは存在するが、企業ごとに提供する情報に差があるため、Python の re ライブラリを用いたルールベース化前処理コードの作成も容易ではない。(連結経営指標等の項目のみであれば可能であるが、それぞれのテキスト項目ごとにコード生成が必要となる)他に考えられるアプローチとしては手作業により正解の教師データを用意し、それを自然言語処理(Sequence-to-Sequence Models など)で学習させることにより適切な前処理変換への手助けとなることが考えられるだろう。

変換後に前述した自然言語処理の前処理プロセスを行うことができるが、⑤の単語分散表現は同じ単語であったとしても状態によってその単語の持つ意味が異なる可能性がある。文章全体の意味や文脈を考慮した上での数値表現を獲得する必要があるが、管理会計に特化した埋め込みレイヤーの開発によって、管理会計への支援精度の向上したモデルの開発へとつながることが推測されるが、開発コストと精度向上による業績に与える影響の費用対効果が合うかは未知数である。

図表 25 テキストの前処理後の CSV

A ^B C Processed_Title	A ^B C Processed_Content
[[0, 17976, 5, 14487, 1362, 28799, 2882, 12, 437, 12662, 2]]	[[0, 28427, 3, 15, 28427, 3, 37, 451, 4, 17976, 4, 744, 3227, 8, 28427, 3...
[[0, 1045, 6982, 166, 134, 24178, 6322, 4, 20298, 25, 28427, 3, 28427, ...	[[0, 5919, 7065, 129, 7, 18037, 8923, 2839, 28427, 3, 32, 5000, 5, 352, ...
[[0, 3596, 12605, 12, 24542, 657, 28427, 3, 28427, 3, 37, 28427, 3, 32, ...	[[0, 2839, 20345, 7, 1989, 8923, 2839, 28427, 3, 32, 5000, 28427, 3, 28...
[[0, 5294, 71, 1764, 129, 543, 9338, 35, 32, 29234, 8, 9435, 3715, 9, 36...	[[0, 962, 28552, 29365, 28528, 4, 5294, 2945, 28973, 543, 9338, 8, 284...
[[0, 1663, 28116, 2938, 7, 3276, 28427, 3, 232, 28775, 28664, 1938, 35...	[[0, 232, 28775, 28664, 4, 24779, 494, 29686, 1938, 8, 89, 962, 1848, 7...
[[0, 28427, 3, 5, 2083, 1800, 151, 35, 28427, 3, 15, 2002, 7, 1652, 36, 2...	[[0, 21409, 28427, 3, 4, 4559, 5, 4190, 1072, 151, 2636, 152, 1252, 286...
[[0, 5738, 2951, 5, 154, 1789, 4346, 51, 3473, 1435, 8540, 7, 12550, 2]]	[[0, 5738, 152, 4, 9498, 1263, 28515, 2951, 8, 89, 962, 1848, 4, 4295, 7...
[[0, 700, 28427, 3, 10648, 3690, 28427, 3, 35, 6981, 3074, 40, 3736, 14...	[[0, 8586, 11568, 8, 2229, 530, 12, 8, 4608, 10, 23, 56, 869, 765, 9, 381...
[[0, 28427, 3, 178, 30347, 2101, 35, 1085, 29431, 4216, 10, 4, 8127, 10...	[[0, 4926, 13747, 28427, 3, 4, 436, 117, 10, 3715, 26, 33, 10, 5, 5581, 1...
[[0, 77, 125, 18164, 996, 28427, 3, 28940, 542, 11618, 28427, 3, 28427...	[[0, 28427, 3, 15, 4, 77, 125, 18164, 996, 11, 28427, 3, 15, 1980, 7, 284...
[[0, 19256, 7474, 4244, 28427, 3, 9894, 1629, 28427, 3, 28427, 3, 2842...	[[0, 57, 29947, 28461, 4, 12766, 51, 41, 1223, 693, 11, 642, 9, 6563, 44...
[[0, 35, 5237, 4, 23398, 36, 8586, 11568, 28427, 3, 700, 52, 11, 24421, ...	[[0, 12285, 4, 120, 2815, 237, 11, 23445, 1657, 6, 35, 5068, 4, 764, 36, ...
[[0, 7094, 4755, 5, 162, 221, 12, 4237, 28427, 3, 686, 4976, 28427, 3, 2...	[[0, 28427, 3, 3710, 28427, 3, 8, 7094, 4755, 352, 6237, 154, 861, 4, 22...
[[0, 7985, 5, 2324, 1165, 2]]	[[0, 2324, 1165, 28427, 3, 32, 328, 7249, 453, 5673, 834, 129, 1518, 49...
[[0, 28427, 3, 15, 20, 52, 4150, 4, 4068, 1613, 8, 28427, 3, 24061, 2842...	[[0, 28427, 3, 15, 9, 4033, 6, 16273, 1961, 17108, 8, 28427, 3, 15, 7, 50...
[[0, 35, 570, 28626, 3530, 36, 89, 55, 28427, 3, 21191, 8, 1680, 4, 15, 7...	[[0, 28427, 3, 15, 11, 327, 457, 6, 21191, 25, 1971, 4068, 172, 4, 4307, ...
[[0, 27071, 8924, 1012, 28586, 5, 449, 28772, 8, 54, 396, 19123, 19, 13...	[[0, 1610, 2932, 5788, 28427, 31997, 16, 29763, 28447, 4, 15, 29383, 5...
[[0, 35, 15257, 28427, 3, 3426, 20588, 36, 28427, 3, 28427, 3, 15, 8, 28...	[[0, 28427, 3, 15, 8, 9600, 1061, 11, 652, 13, 5, 771, 9, 9172, 894, 4, 28...
[[0, 35, 826, 101, 825, 36, 28427, 3, 2001, 17399, 172, 7, 7346, 7737, 2]]	[[0, 1148, 121, 17399, 172, 11, 28427, 29597, 2588, 4, 3189, 9, 8602, 3...
[[0, 28427, 3, 28427, 3, 28427, 3, 17399, 172, 5, 15197, 387, 28899, 35...	[[0, 28427, 3, 15, 4, 15, 11, 327, 305, 25, 28427, 3, 566, 551, 4, 35, 284...
[[0, 35, 3136, 2219, 36, 4, 59, 28450, 28527, 28427, 3, 21192, 360, 293...	[[0, 702, 888, 11, 15769, 35, 172, 29046, 28575, 275, 36, 6, 988, 7, 320...
[[0, 4627, 3985, 101, 6393, 28441, 12389, 28722, 426, 7, 2]]	[[0, 28427, 3, 15, 4, 1438, 28579, 28500, 10, 728, 1329, 592, 8, 1078, 4...

出典：筆者作成

参考にテキストをテンソル変換したデータは図表 25 の通りであり、このようなテキストが数値化されたデータを用いて LLM の学習を行うこととなる。

6.4 学習

本論文における学習プロセスは事前学習済み LLM を用いて、ファインチューニングを行い予算管理に特化したモデルを作成することにある。ライブラリは LLMLTime を用いることで財務会計上の数値データをテキストデータで補完し、予測に役立てることとなる。

学習時における問題は計算リソースの問題である。LLM の学習は AI の中でも多くの計算リソースを必要とし、「NVIDIA A100」などの高性能かつ高価格の GPU を複数用意しなければならないケースもある。演算コストを抑える技術である量子化などの技術(Bai et al.2020)の研究も進みつつあるが、現状においては LLM の学習環境を構築するには金銭的な障壁が存在する。多くの人工知能の研究を行う研究室では AI 用の研究費の確保がなされているが、会計の研究では AI に特化した GPU は用意することが困難である。

代替的手法として、AWS などのクラウドサービスを通じて計算リソースを借りるという手段があるが、欠点として定量課金制であるため試行錯誤を繰り返す AI の開発には結果として多くの費用がかかる可能性がある点である。

大規模なモデルの利用(70b パラメーターなど)では多くの計算リソースを必要とするが、規模が小さめのモデル(7b パラメーターなど)では一般向けの GPU でも学習が可能である。しかし、規模の小さいモデルでは当然ながら学習データが少ないため望んだ応答が得られることはまずないといって良い。規模の小さいモデルでは次単語の予測がうまく行えず入力したプロンプトを繰り返すことも多い。

(例:プロンプト/日本で一番高い山はなんですか？

応答/ 日本で一番高い山はなんですか？日本で一番高い山はなんですか？日本で一番高い山はなんですか？…)

よって管理会計タスクをこなせる LLM の学習のためには規模の大きいモデルで開発することが求められるが、環境が用意できない場合であれば計算コストの低い規模の小さいモデルでの検証を行うことで、試行錯誤が可能となりコストを抑えることができる。

6.5 チューニング等の調整

予算管理 LLM では一般的なハイパーパラメーターチューニングに加えて、RLHF によって望んだ回答を作成することが要求される。特に予算管理に特化したモデルを生成する場合には望まれる回答フォーマットに準じた応答が求められるため、多くのフィードバックを用意する必要がある。

Ouyang et al.(2022)は言語モデルにおける強化学習の問題点として「Reward Hacking」と「Alignment Tax」を挙げており調整時には注意が必要である。

Reward Hacking とは強化学習が報酬に基づいて価値を決定してしまうことから、報酬が最大化されるような文ばかり生成してしまうようになり、生成に柔軟性が欠けてしまうことである。

Alignment Tax とは人間が意図する通りに事前学習モデルを学習させると知識の忘却が起こり、性能が劣化することである。

これらの問題点を防ぐために KL Penalty や Replay といった項目の調整が必要となる。

6.6 完成後の調整・評価

AI システムは従来のパッケージ型ソフトウェアなどとは違い、実装後も継続的に開発・監視・調整を続けていく必要がある(Steidl et al.2023)。管理会計タスクにおける AI は精度の検証が困難である。例えば株価予測 AI(Kamalov et al.2020)では予測値と実際値の誤差によって精度の検証を可能とするが、管理会計においてそういった素早く精度を図る指標はなく、企業の年間の活動後のフィードバックを通じて評価を行う必要がある。

予算編成 AI のうち「Base Model」は財務諸表の予測数値と実際の数値を比較することによって可視性の高い精度検証が可能である。ただ、Base Model の目的は財務諸表数値を一致させることではなく、管理会計タスクに用いるための将来財務諸表数値を合理的な水準で算出することを目的としたモデルであるため、単純に誤差を比較するのではなく管理会計タスクのアシストを行える数値を算出できているかどうかで評価する必要がある。

「予算管理 Model」ではより精度の検証が難しくなる。業務代替的な AI であれば人件費の削減や作業の効率化といったコスト削減の尺度を持った AI の評価が行えるが、予算管理業務の支援を行う予算管理 Model では予算の適正化による予算に対する満足度の向

上や環境，経済，市場を踏まえた上での予算管理による業績向上などによる評価が考えられる。これらの評価は単年度で判断することは難しく，長期的なスパンでの評価が求められる。

第7章 おわりに

7.1 本研究の貢献

本論文では管理会計・予算管理における AI の適用可能性について検討し、実態調査の整理を通じて得られた問題点を解決へ導くための予算管理 LLM の構想を行なった。また、実際の AI 構築時に発生する障壁について明らかとし、今後の管理会計の AI 開発に向けて展望を示した。

管理会計の AI 利用を検討した先行研究ではアルゴリズムやライブラリにまで言及した検討を行なっているケースは少なく、AI を適用することで予測される影響を論じているものの、具体的な手法について明記していない先行研究(Ulrich et al.2022;Vărzaru 2022; Losbichler and Lehner 2021)が見られた。そこで本研究では実態調査を通じて企業が予算管理において課題だと感じている項目について具体的な AI での解決手段の検討を行うことで今後の管理会計の AI 活用の指針を示しながら、予算編成 AI のあり方を探求した。

多くの企業が課題であると感じている予算管理に掛かる期間やコストを AI によって対応することは、現状では難しかったが、その他の問題点である環境変化への対応や、予算戦略とのリンク性、ストレッチな予算の適用、予算実績差異分析といった項目は、それぞれ AI によってアプローチの可能性があることを論じた。これらの問題は予算管理 LLM によってアプローチを行うことで AI 化によって、予算の持つ諸機能が失われずに予算管理の支援ができるとし、予算管理 LLM の構想と必要となる AI の技術について管理会計的視点から検討を行なった。

予算管理 LLM は Gruver et.al(2023)の LLMTime を用いることでデータが少なく時系列性を持つ財務諸表数値を経済、ビジネス実態を反映するテキスト情報で補完し、時系列の特性を残したまま予測を行うアプローチによって財務諸表を合理的な水準で予測するモデルを提案した。

また実際に AI システム構築過程を得て経験した管理会計 AI モデル構築上の障壁の整理を行い、現在黎明期であり導入が進んでいない管理会計 AI 開発時に起こりうる問題点を明らかにした。

今回明らかとなった障壁の例としては、企業は管理会計 AI 導入のためのインフラが整

っていない可能性があり、管理会計 AI を推進していくためのデータの入手性に難があることや、有価証券報告書の XBRL の形式として財務数値は構造化されており、データ分析に活用しやすい形式で提供されているもののテキストデータは分析に向かない形式での取得となること、経済状況を把握するためのニュースデータは多くの提供会社が AI の学習への利用に制限をかけていること、LLM は計算コストが非常に高く専門的な設備を有する環境でなければ開発を行うことが難しいことなどが明らかとなった。

7.2 本研究に残された課題

本研究では実際に予算編成 AI システムの開発を試みることで、予算編成 AI 開発上の障壁を明らかとしたが、時間的・知識的・資源的制約によりシステムを完成させ、実際の精度検証や効果検証を行うことができなかった。また、システムを運用する上での問題点や調整時に生じる課題点なども存在すると想定されることから、今後の管理会計導入の指針としては物足りないものとなった。

AI 開発では教師データ収集上の制約や開発環境上の制約が発生することから、全ての要望を満たした上での開発を行うことは難しく、経済的問題も考慮した上で妥協点を探し出し限られた資源で開発を行う必要がある。

実用レベルの精度で支援を行う管理会計 AI の開発には多くの企業の協力が必要不可欠であり、その環境を作り上げるためには限られた資源の中で、管理会計に AI を利用することで得られる便益があることを証明する必要があるだろう。

そこで、入手性の高いデータや環境を用いて、予算管理 LLM の Base Model の完成・調整・検証をすることで、将来的に管理会計 AI を積極活用した際に得られるメリットの可能性を検討し、今後の管理会計 AI 活用に価値を見出すことができるかの検証を今後の課題とする。

謝辞

本論文の制作にあたり、指導教員の皆様や研究室の方々からの執筆ならびに AI 開発の支援によって本論文の完成に至りました。

主査である目時壮浩先生には AI という特殊な研究であったにもかかわらず、多大なる支援と貴重な指導を賜りましたことを厚く御礼申し上げます。

副査である鈴木孝則先生、長谷川恵一先生には熱心な指導と終始適切な助言に感謝申し上げます。また、早稲田大学大学院商学研究科博士後期課程 王聖書様には長期にわたる研究指導を頂きましたこと感謝の意を申し上げます。

1 年半の間、切磋琢磨し論文を書き終えた仲間である伊藤立基様、川原悠様、工藤桂菜様、長岡将史様、渡辺理奈様にも感謝の述べさせていただきます。

参考文献

- 足立洋.2017.「現場サポート」の予算管理:中小企業における予算管理と人的コミュニケーションの関係性」『中小企業会計研究』 3:25-36.
- 安立竜清・上向井麻希・得重侑弥・前田ひなた・村澤莉依・行武香音.2022.「組織の人事評価における AI フィードバックと個別的信頼の形成」『人工知能学会全国大会論文集』 36:1-4.
- 伊藤克容.2007.「日本企業の予算管理システムの特徴と稟議制度の影響」『成蹊大学経済学部論集』 38(1):17-39.
- 伊藤嘉博・目時壮浩.2021.『異論・正論管理会計』中央経済社.
- 妹尾剛好.2017.「日本企業の予算管理の類型と探索・深化との関連分析:探索的研究」『原価計算研究』 41(1):38-50.
- 妹尾剛好・横田絵理.2013.「日本企業における予算に基づく業績評価に関する考察:主観的評価に焦点をあてて」『原価計算研究』 37(1):96-106.
- 上東正和.2021.「わが国企業の企業規模と管理会計実践の実態」『富山大学経済論集』 67(1):89-113.
- 上山晋平.2021.「予算管理上の稟議制度の機能に関する実態調査:稟議による責任及びアカウンタビリティ共有機能について」『メルコ管理会計研究』 13(1):19-28.
- 鬼塚雄大.2018.「本社による在外子会社の業績管理:業績管理システム運用への着目」『原価計算研究』 42(1):58-70.
- 川野克典.2014.「日本企業の管理会計・原価計算の現状と課題」『商学研究』 30:55-86.
- 企業予算制度研究委員会編.2018.『日本企業の予算管理の実態』中央経済社.
- 岸田隆行.2013.「日本企業における予算管理実務:質問票調査の結果報告」『駒大経営研究』 44(1-2):21-45.
- 岸田隆行.2020.「適正な予算目標と予算スラック」『駒大経営研究』 51(1-2):171-184.
- 栗原聡・山川宏.2015.「人工知能とは(第12回)」『journal of the Japanese Society for Artificial Intelligence,』 30 (1):104-110.
- 清水孝.2016.「わが国における予算管理実務改善に関する調査」『早稲田商学』 446:103-130.

- 清水孝・町田遼太・上田巧.2019.「我が国における予算管理の改善に関する研究の動向:脱予算経営の観点から」『早稲田商学』455:1-31.
- 清水孝・町田遼太・上田巧.2021.「日本企業の予算管理の改善に関する実態調査」『早稲田商学』460:207-244.
- 菅本栄造・宮脇貴代之.2020.「公共土木一式工事会社の先進的な実行予算管理システムの研究:ミヤシステム(株)および導入会社に対するインタビュー調査にもとづいて」『青山経営論集』55(2-3):3-32.
- 鈴木章悟・小澤誠一・渡辺一男・廣瀬勇秀・池田佳弘・飯塚正昭・西田大輔.2022.「機械学習を用いたアナリストレポート分析と投資判断レーティング予測」人工知能学会第二種研究会資料 28:164-171.
- 谷守正行.2019a.「管理会計への AI 適用可能性」『企業会計』71(2):29-36.
- 谷守正行.2019b.「AI 管理会計に関する理論的研究:将来予測とフィードフォワードへの AI 適用可能性」『会計学研究』45:75-101.
- 谷守正行.2020.「AI 意思決定会計の研究 : AI による行動経済学と意思決定の統合化」『専修商学論集』110:135-146.
- 谷守正行.2021.「原価計算への AI 適用研究:銀行 ABC の課題解決の可能性」『専修商学論集』112:97-108.
- 谷守正行.2023.「AI による管理会計の進化と変革」『証券アナリストジャーナル』61(9):16-25.
- 飛田努.2012.「中小企業における経営管理・管理会計実践に関する実態調査:福岡市内の中小企業を調査対象として」『会計専門職紀要』3:57-69.
- 中村彰良.2015.「脱予算管理と日本企業への適用可能性」『高崎経済大学論集』57(4):17-28.
- 西信洋・金田直之.2009.「経営者予想の信頼性」『学習院大学 経済論集』45(4):269-292.
- 日本経済新聞.2023.『米 NY タイムズ、OpenAI を提訴 記事流用で数千億円損害』日本経済新聞社 (<https://www.nikkei.com/article/DGXZQOGN27CXP0X21C23A2000000/> 閲覧日:2024年1月8日).
- 早川翔・妹尾剛好・新井康平・安酸建二・横田絵理.2018.「予算期間と予算修正の方法が財務業績に与える影響:探索的研究」『原価計算研究』42(2):67-78.
- 広原雄二.2021.「管理会計の適合性について」『東北文化学園大学総合政策学部紀要』20(1):19-32.

- 堀井悟志.2015.「予算管理とイノベーションの創出」『管理会計学』23(1):61-71.
- 堀井悟志・早川正明・横田明紀.2021.「SFA と連携した BI 活用によるフィードフォワード型予算管理の実現」『原価計算研究』45(2):26-38.
- 榎谷奎太・岩澤佳太・吉田栄介.2022.「日本企業における業績管理の変化と変容:10年分の実態調査データに基づく分析と考察」『メルコ管理会計研究』13(2):3-20.
- 山田庫平・鈴木研一・山下裕企・大槻晴海・三木僚佑「わが国企業予算制度の実態(平成14年度)(2)企業予算制度の基礎的事項に関する分析:予算編成目的,経営計画,予算委員会,予算期間等」『産業経理』63(2):120-135.
- 横田絵理・妹尾剛好.2011.「予算管理への影響要因:予算編成・目標の困難化と戦略リンクの強化への影響分析」『原価計算研究』35:107-119.
- 吉田栄介・岩澤佳太・徐智銘・榎谷奎太.2019.「日本企業における管理会計の実態調査(第3回)業績・予算管理編:東証・名証1部上場企業」『企業会計』71(11):1565-1571.
- Al-Htaybat, K., and L. von Alberti-Alhtaybat. 2017. Big Data and corporate reporting: impacts and paradoxes. *Accounting, Auditing & Accountability Journal* 30 (4):850-873.
- Amato, F., F. Guignard, S. Robert, and M. Kanevski. 2020. A novel framework for spatio-temporal prediction of environmental data using deep learning. *Scientific Reports* 10 (1):1-11.
- Bahrini, A., M. Khamoshifar, H. Abbasimehr, R. J. Riggs, M. Esmaceli, R. M. Majdabadkohne, and M. Pasehvar. 2023. ChatGPT: Applications, Opportunities, and Threats. *2023 Systems and Information Engineering Design Symposium: 274-279.*
- Bai, H., W. Zhang, L. Hou, L. Shang, J. Jin, X. Jiang, Q. Liu, M. Lyu, and I. King. 2020. BinaryBERT: Pushing the Limit of BERT Quantization. *arXiv.org*:2012.15701.
- Bhimani, A., and L. Willcocks. 2014. Digitisation, 'Big Data' and the transformation of accounting information. *Accounting and Business Research* 44 (4):469-490.
- Brown, T. B., B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei. 2020. Language Models are Few-Shot Learners. *Neural Information Processing Systems 2020*:1-72.

- Casas-Arce, P., S. M. Lourenço, and F. A. Martínez-Jerez. 2017. The Performance Effect of Feedback Frequency and Detail: Evidence from a Field Experiment in Customer Satisfaction. *Journal of Accounting Research* 55:1051-1088.
- Chen, D., J. P. Esperança, and S. Wang. 2022. The Impact of Artificial Intelligence on Firm Performance: An Application of the Resource-Based View to e-Commerce Firms. *Frontiers in Psychology* 13:1-14.
- Christiano, P., J. Leike, T. B. Brown, M. Martic, S. Legg, and D. Amodei. 2023. Deep Reinforcement Learning from Human Preferences. *arXiv:1706.03741*.
- Church, K. W., Z. Chen, and Y. Ma. 2021. Emerging trends: A gentle introduction to fine-tuning. *Natural Language Engineering* 27 (6): 763-778.
- Cun, Y. L., B. Boser, J. S. Denker, R. E. Howard, W. Hubbard, L. D. Jackel, and D. Henderson. 1990. Handwritten digit recognition with a back-propagation network. *Natural Information Processing Systems* 2:396-404.
- Dahlan, M. 2019. Analysis of interrelationship between usefulness of management accounting systems, interactive budget use and job performance. *Management Science Letters* 9 (7):967-972.
- Dai, Q. 2022. Designing an Accounting Information Management System Using Big Data and Cloud Technology. *Scientific Programming* 2022:1-11.
- Dai, Z., Z. Yang, Y. Yang, J. Carbonell, Q. V. Le, and R. Salakhutdinov. 2020. Transformer-XL: Attentive Language Models Beyond a Fixed-Length Context. *Proceedings of the Conference*:2978-2988.
- Dent, J. F. 1990. Strategy, organization and control: Some possibilities for accounting research. *Accounting Organizations and Society* 15 (1):3-25.
- Elbeltagi, E., and H. Wefki. 2021. Predicting energy consumption for residential buildings using ANN through parametric modeling. *Energy Reports* 7:2534-2545.
- Frey, C. B., and M. A. Osborne. 2017. The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting & Social Change* 114:254-280.
- Fu, D. Y., T. Dao, K. K. Saab, A. W. Thomas, A. Rudra, and C. Ré. 2022. Hungry Hungry Hippos: Towards Language Modeling with State Space Models. *arXiv:2212.14052*.
- Glazova, M. V., O. M. Korobeynikova, and E. V. Yagupova. 2022. Digital transformation of

- processes in the management accounting system. *Series Economics and Law* 32 (3):432-437.
- Granstrand, O., and C. Oskarsson. 1994. Technology diversification in "MUL-TECH" corporations. *IEEE Transactions on Engineering Management* 41 (4):355-364.
- Czum, J. M. 2020. Dive Into Deep Learning. *Journal of the American College of Radiology* 17 (5): 637-638.
- Gärtner, B., and M. R. W. Hiebl. 2017. *Issues with Big Data*. (<https://www.taylorfrancis.com/chapters/edit/10.4324/9781315647210-13/issues-big-data-bernhard-g%C3%A4rtner-martin-hiebl>. 最終閲覧日:2023年12月30日).
- Gruver, N., M. Finzi, S. Qiu, and A. G. Wilson. 2023. Large Language Models Are Zero-Shot Time Series Forecasters. *arXiv:2310.07820*
- Guo, M., J. Ainslie, D. Uthus, S. Ontanon, J. Ni, Y.H. Sung, and Y. Yang. 2022. LongT5: Efficient Text-To-Text Transformer for Long Sequences. *Cornell University Library*:724-736.
- He, K., X. Zhang, S. Ren, and J. Sun. 2016. Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition*:770-778.
- Hochreiter, S., and J. Schmidhuber. 1997. Long Short-term Memory. *Neural Computation* 9 (8):1735-1780.
- Hope, J., and R. Fraser. 2003. Who Needs Budget. *Harvard Business Review* 81:108-115.
- Jin, M., S. Wang, L. Ma, Z. Chu, J. Y. Zhang, X. Shi, P. Y. Chen, Y. Liang, Y. F. Li, S. Pan, and Q. Wen. 2023. TIME-LLM: Time Series Forecasting by Reprogramming Large Language Models. *arXiv:2310.01728*.
- Kamalov, F., L. Smail, and I. Gurrib. 2020. Stock price forecast with deep learning. *In Proceedings of the 2020 International Conference on Decision Aid Sciences and Application* :1098-1102.
- Kaplan, R. S., and D. P. Norton. 1996. *The Balanced Scorecard: Translating Strategy into Action*. untag-smd. (http://www.untag-smd.ac.id/files/Perpustakaan_Digital_1/BALANCED%20SCORECARD%20The%20balanced%20scorecard%20translating%20strategy%20into%20action%20%5B1996%5D.pdf. 最終閲覧日:2023年9月27日).

- Kokina, J., and S. Blanchette. 2019. Early evidence of digital labor in accounting: Innovation with Robotic Process Automation. *International Journal of Accounting Information Systems*35:1-13.
- Korhonen, T., E. Selos, T. Laine, and P. Suomala. 2021. Exploring the Programmability of Management Accounting Work for Increasing Automation: An Interventionist Case Study. *Accounting auditing & accountability journal*34(2):253-280.
- Krahel, J., and W. Titera. 2015. Consequences of Big Data and formalization on accounting and auditing standards. *Accounting Horizons* 29 (2):409-422.
- Krizhevsky, A., I. Sutskever, and G. E. Hinton. 2012. ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems 2012* (2):1097-1105.
- Li, C., J. Wang, Y. Zhang, K. Zhu, W. Hou, J. Lian, F. Luo, Q. Yang, and X. Xie. 2023. Large Language Models Understand and Can be Enhanced by Emotional Stimuli. *arXiv:2307.11760*.
- Libby, T., and R. M. Lindsay. 2010. Beyond Budgeting or Budgeting Reconsidered? A Survey of North-American Budgeting Practice, *Management Accounting Research* 21(1):56-75.
- Lightman, H., V. Kosaraju, Y. Burda, H. Edwards, B. Baker, T. Lee, J. Leike, J. Schulman, I. Sutskever, and K. Cobbe. 2023. Let's Verify Step by Step. *arXiv:2305.20050*.
- Losbichler, H., and O. M. Lehner. 2021 Limits of Artificial Intelligence in Controlling and the Ways Forward:a Call for Future Accounting Research. *Applied Accounting Research* 22 (2):365-382.
- Marotta, G., and C. D. Au. 2021. *Opportunities of AI in Budgeting Approaches*. Digitale Welt. (https://www.researchgate.net/publication/359323666_Opportunities_of_AI_in_budgeting_approaches. 最終閲覧日:2023年9月27日).
- Marotta, G., P. Krahhof, C. D. Au. 2022. Bedeutung von Verhandlungstechniken im Rahmen der Budgetplanung. *Social Science Research Network*:1-13.
- Matsumoto, Y., A. Suge, and H. Takahashi. 2023. Analysis of the relationship between technological diversification and enterprise value using patent data. *Information*

Technology and Management:1-24.

- Moffitt, K., and M. Vasarhelyi. 2013. AIS in an age of Big Data. *Journal of Information Systems* 27 (2):1-19.
- Namir, K., H. Labriji, and E. H. Ben Lahmar. 2022. Decision Support Tool for Dynamic Inventory Management using Machine Learning, Time Series and Combinatorial Optimization. *Procedia Computer Science* 198:423-428.
- Neely, A., M. Bourne, and C. Adams. 2003. Better Budgeting or Beyond Budgeting?. *Business Excellence* 7(3):22-28.
- Nishi, Y., A. Suge, and H. Takahashi. 2021. Construction of a news article evaluation model utilizing high-frequency data and a large-scale language generation model. *SN Business & Economics* 1 (8):1-18.
- Ouyang, L., J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, J. Schulman, J. Hilton, F. Kelton, L. Miller, M. Simens, A. Askell, P. Welinder, P. Christiano, J. Leike, and R. Lowe. 2022. Training language models to follow instructions with human feedback. *arXiv:2203.02155*
- PricewaterhouseCoopers. 2023. *AI Business Predictions*. PricewaterhouseCoopers(<https://www.pwc.com/us/en/tech-effect/ai-analytics/ai-predictions.html>.最終閱覽日:2023 年 11 月 5 日).
- Radford, A., J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever. 2019. *Language Models are Unsupervised Multitask Learners*. Papers With Code.(<https://paperswithcode.com/paper/language-models-are-unsupervised-multitask>. 最終閱覽日:2023 年 1 月 25 日).
- Rumelhart, D. E., Hinton, E. Geoffrey, and J. Ronald. 1985. Learning Internal Representations by Error Propagation. *ICS Reports* 8506:1-49.
- Rumelhart, D. E., G. E. Hinton, and R. J. Williams. 1986. Learning representations by back-propagating errors. *Nature* 323 (6088):533-536.
- Seyedan, M., and F. Mafakheri. 2020. Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. *Journal of Big Data* 7 (1):1-22.
- Sponem, S., and C. Lambert. 2016. Exploring differences in budget characteristics, roles and

- satisfaction: A configurational approach. *IDEAS Working Paper Series from RePEc* 30:47-61.
- Steidl, M., M. Felderer, and R. Ramler. 2023. The pipeline for the continuous development of artificial intelligence models—Current state of research and practice. *The Journal of Systems and Software* 199:1-26.
- Suneja, N., J. P. Shah, Z. H. Shah, and M. S. Holia. 2021. A neural network approach to design reality oriented cost estimate model for infrastructure projects. *Reliability--Theory & Applications* 16:254-263.
- Thomas Publishing Company.2019. *Big Data, Better Budgeting: Machine Learning for Facilities Management*.(<https://www.thomasnet.com/insights/big-data-better-budgeting-machine-learning-for-facilities-management/>.最終閱覽日:2023年12月28日).
- Ulrich, P., V. Frank, R. Buettner, and W. Becker. 2022. A Literature Review on the Impact of Modern Technologies on Management Reporting. *Procedia Computer Science* 207:907–915.
- Varzaru, A. A. 2022. Assessing Artificial Intelligence Technology Acceptance in Managerial Accounting. *Electronics (Basel)* 11 (14):1-14.
- Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. 2017. Attention Is All You Need. *Neural Information Processing Systems* 2017:5999-6009.
- Wahyuningsih, D., D. Caisar Darma, R. A. Nuraliaty, J. Kasuma, T. Tasente, S. Maria, D. Ushakov, and Sriwardani.2020. Why Dynamic Capacity Influences the Quality of Management Accounting Information Systems in the Public Sector?. *International Journal of Psychosocial Rehabilitation*24(10):4032-4044.
- Wang, P. 2019. On Defining Artificial Intelligence. *Journal of Artificial General Intelligence* 10(2): 1-37.
- Wei, J., X. Wang, D. Schuurmans, M. Bosma, B. Ichter, F. Xia, E. H. Chi, Q. V. Le, and D. Zhou. 2022. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. *Neural Information Processing Systems*35:1-43.
- Warren, J. D., K. C. Moffitt, and P. Byrnes. 2015. How big data will change accounting. *Accounting Horizons* 29 (2):397-407.

- Wu, S., O. Irsoy, S. Lu, V. Dabrovolski, M. Dredze, S. Gehrmann, P. Kambadur, D. Rosenberg, and G. Mann. 2023. BloombergGPT: A Large Language Model for Finance. *arXiv:2303.17564*.
- Wu, X., and S. Du. 2022. An Analysis on Financial Statement Fraud Detection for Chinese Listed Companies Using Deep Learning. *IEEE Access* 10: 22516-22532.
- Zaheer, M., G. Guruganesh, A. Dubey, J. Ainslie, C. Alberti, S. Ontanon, P. Pham, A. Ravula, Q. Wang, L. Yang, and A. Ahmed. 2020. Big Bird: Transformers for Longer Sequences. *arXiv:2007.14062*.
- Zemánková, A. 2022. Artificial intelligence in management accounting. *Český Finanční a Účetní Časopis* (4):81-99.
- Zhang, Y., K. Gong, K. Zhang, H. Li, Q. Yu, W. Ouyang, and X. Yue. 2023. Meta-Transformer: A Unified Framework for Multimodal Learning. *arXiv: 2307.10802*.
- Zhang, P., X. Dai, J. Yang, B. Xiao, L. Yuan, L. Zhang, and J. Gao. 2021. Multi-Scale Vision Longformer: A New Vision Transformer for High-Resolution Image Encoding. *IEEE/CVF International Conference on Computer Vision (ICCV):2978-2988*.
- Zhang, X. 2021. Application of data mining and machine learning in management accounting information system. *淡江理工學刊* 24 (5):813-820.
- Zhao, Y., G. Wang, C. Tang, C. Luo, W. Zeng, and Z. J. Zha. 2021. A Battle of Network Structures: An Empirical Study of CNN, Transformer, and MLP. *arXiv: 2108.13002*.
- Zott, C., R. Amit, and L. Massa. 2011. The Business Model: Recent Developments and Future Research. *Journal of management* 37(4):1019-1042.