

# 生成AIを活用して進める「モデルリスク管理」の効率化

## 管理モデル数の増大を見据えて レポート自動作成の可能性を探れ

**金** 融機関において「モデル」の活用が広がっている。最近では、マネー・ローンダリング対策や不正検知等のコンプライアンス、気候変動といった新たな領域でもモデルが活用され、管理すべきモデル数が増えている。一方で、モデルの管理主体である2線のリソース確保は、特に難しい状況にある。本稿では、負担の高まるモデルリスク管理において、高度な文章作成能力を持つ生成AIを活用した検証レポートの自動作成の可能性を示し、金融機関（特に2線）の業務効率化策を提案する。

### モデルリスク管理を巡る 負担増大とリソース不足

金融機関では多くの「モデル」が業務に使用されていることを受けて、モデルリスク管理（Model Risk Management ≡ MRM）に対する関心が高まっていく。MRMを巡っては、米国

が2011年に、モデル開発から検証、承認、モニタリング、再検証、使用停止までのモデル・ライフサイクルに応じたガバナンスに関するガイダンス（注1）を公表した。本邦では、21年に金融庁が「モデル・リスク管理に関する原則」（以下、金融庁原則。注2）を公表。英国

もガイダンス（注3）を23年に公表するなど、各国が相次いでガイダンスを公表している。MRMでは、開発したモデルに対して、1〜3線それぞれによる管理が求められる。具体的には、①モデルのオーナーや使用者等（1線）によるモニタリング、②リスク管理部門（2

線）によるモデルリスクの管理とモデル（再）検証、③内部監査部門（3線）によるMRM態勢の監査——である。金融庁原則の対象となる本邦大手金融機関では、保有するモデル数が数百以上、米国GIBS（グローバルなシステム上重要な銀行）であれば数千以

あずさ監査法人

金融統轄事業部

金融アドバイザリー事業部

シニアマネジャー

竹川 正浩

ディレクター

田中 康浩

マネージングディレクター

曾我部 淳

上に達することがある。最近では、マネー・ローディング対策や不正検知等のコンプライアンス、気候変動といった新たな領域でもモデルが活用されている。管理すべきモデル数は今後とも間違いなく増加するだろう。

MRMは、2線において、特にモデル検証や検証リポートの作成の負担が重いことで知られている。モデルのリスクにもよるが、検証リポートは、数十から100ページ程度のボリュームになることが通常である。

一方、本邦では大手金融機関であつても、モデルやMRMに知見のある2線のリソースが不足している。また、現在は金融庁原則の適用対象外の金融機関でも、今後MRMの態勢構築が求められる可能性もあるが、こうした金融機関ではさらにリソースが不足している。以上のような状況を踏まえると、2線のモデル検証やリポート作成の業務を効率化し、負担を軽減することは重要である。

## 試行実験における 検証リポートの自動作成

足元では生成AIの活用が広がっている。23年はChat（チャット）GPTに代表される生成AIの認知が一気に進み、企業が盛んに導入を進めた。金融機関も、翻訳や社内Q&Aですでに利用しており、今後もコールドセクターやコンプライアンス領域の自動化・効率化といった、より高度な利用を探る動きが継続しそうだ。

生成AIは、高い文章作成能力を有することに加えて、検証プログラムの自動作成やタスクの網羅的なリストアップに長ける。この点、膨大な検証リポートの作成が求められるMRMの特徴に着目し、金融機関の業務効率化に向けて生成AIの活用による検証リポート自動作成の可能性を探りたい。

本稿で紹介する試行実験では、金融機関のストレステスト等で使用するPD（デフォルト確率）推計モデルを例に取り上げる。当該モデルは、リスク管理部門等において、年次で精度検証を行うケースが多く、検証手法は二項検定（実績デフォルト率がPD推計モデルの予測誤差

### ベースとなるプロンプトの主要部分

〔図表1〕

あなたは金融機関におけるモデルリスク管理の専門家です。以下はPD推計モデルの概要とバックテスト結果です。与えられたデータに基づき日本語で検証リポートを作成してください。作成に当たっては、バックテストの数値を単に繰り返すのではなく、バックテストの評価と結論に重点を置いたリポートを作成してください。

### PD推計モデルの概要 ###  
用途：ストレステストにおいてマクロ経済変数からPDを推計するモデル  
被説明変数：PD推計値  
説明変数と回帰係数：(省略)  
モデル構造：線形回帰モデル

### バックテストの手法 ###  
PDに関する二項検定

### バックテスト結果 ###  
(見やすさを優先し下表にまとめているが、実際はプロンプトに直接入力)

時点	期初件数	デフォルト件数	実績デフォルト率	PD推計値	95%信頼区間下限値	95%信頼区間上限値	判定結果
2022/9	6,195	93	1.50%	1.20%	0.94%	1.49%	NG
21/9	6,072	81	1.33%	1.19%	0.92%	1.47%	OK
20/9	6,263	70	1.12%	1.22%	0.96%	1.50%	OK

(注) 仮想の数字を用いている。  
(出所) 筆者作成 (図表2～4も同じ)

の範囲内であるかを評価)が一般的に用いられる。試行実験では、仮想的な二項検定の結果をインプットデータとして、生成AIにモデル検証のリポートを自動作成させることを試みる。第1ステップとして、ベース

となるプロンプト(図表1)を準備し、リポートの自動作成を行った。プロンプトは次の4点を意識して作成した(図表1の下線に対応)。  
①文脈 金融機関におけるモデルリスク管理の専門家という役

割を付与

②命令II検証リポートの作成を指示

③出力形式II検証の評価と結論に重点を置いた出力を指示

④入力データII検証結果を認識しやすいように区切り文字#を使用

生成AIは、二項検定の原理を理解しているわけではないため、必ずしも正確な評価および文章が作成されるとは限らない。そこで、二項検定の評価ができる限り正確に出力されるように、入力と出力の例をいくつか与える(Few-shotプロンププディング)など、プロンププの作りを別途工夫している。

## モデル評価基準を模したプロンププの追加

図表2の生成AIの回答(リポートの自動作成)を見ると、二項検定の評価自体は正しい。しかし、「評価と結論」はプロンププの情報が文章に変換された程度にとどまり、物足りない印象を受ける。

この結果を受けて、生成AIとの対話を繰り返すことで記載

【図表2】 ベースとなるプロンププに対する生成AIの回答

<p>バックテストのリポート</p> <ol style="list-style-type: none"> <li>前提条件 バックテストの種類：PDに関する二項検定 PD推計モデルの概要：線形回帰モデルによるPDの推計。説明変数は(省略)</li> <li>2022年9月期のバックテスト結果 22年9月期の実績デフォルト率は1.50%であり、モデルの推計値は1.20%となりました。PD推計値の95%信頼区間の下限値は0.94%、上限値は1.49%です。バックテストの結果、実績デフォルト率は推計値の95%信頼区間の範囲外にあり、判定結果はNGとなりました。従って、モデルは22年9月期のデフォルト予測において十分な精度を持っていないと結論付けられます。</li> <li>21年9月期のバックテスト結果 (22年9月期と類似するため省略)</li> <li>20年9月期のバックテスト結果 (22年9月期と類似するため省略)</li> <li>評価と結論 22年9月期のバックテスト結果を総合的に評価すると、モデルはデフォルト予測において十分な精度を持っていないことが示されました。</li> </ol>
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

内容をより充実させる方法も考えられる。他方、今回の試行実験では、次のステップとしてMRM実務の観点から、モデルリスク管理規程などに定められている評価基準を追加することで、結果がどのように変化するかを確認した。

モデルの性能評価指標や評価基準値を下回った際に取るべきアクションを、モデルリスク管理規程などに定めるケースは多い。ここでは、そのような規程を模したプロンププ(図表3)を追加することで、作成されるリポートの変化を見ていく。

【図表3】 モデルリスク管理規程を模したプロンププ

<p>以下はバックテストの評価基準です。この評価基準に基づき、モデルの継続利用の可否を判定し、その判定結果を結論に含めるかたちで検証リポートを作成してください。</p> <p>###評価基準###</p> <p>判定結果が直近1期のみNGの場合：ストレステスト全体に与える影響を考慮して、モデルの再構築の要否を検討する。</p> <p>判定結果が直近2期連続でNGとなった場合：モデルの再構築を実施する。</p>
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

【図表4】 モデルリスク管理規程を踏まえた生成AIの回答

<p>【バックテストのリポート】(1.~4.は図表2と同様のため省略)</p> <ol style="list-style-type: none"> <li>評価と結論 2022年9月期では、モデルの判定結果はNGであり、直近1期のみでNGとなりました。ストレステスト全体に与える影響を考慮し、モデルの再構築の要否を検討する必要があります。ただし、過去のバックテスト結果では、21年9月期と20年9月期においてモデルの判定結果はOKとなりました。 総合的に判断すると、モデルは一部の期において精度に課題がある可能性が示されましたが、過去のバックテスト結果では適切な精度を持っていたことも考慮し、適宜の改善や検証を行いながら継続利用することができると結論付けられます。</li> </ol>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

モデル検証では、直近の検証結果のみで性能を評価することは少なく、複数年度にわたる検証結果を踏まえて評価することが一般的である。ここでも、直近1期のみで検証結果がNGの場合にはモデルの再構築を「要検証」、2期連続NGとなつて初

めて再構築を「実施」という基準を設けている。

これを受けた生成AIの回答（図表4）を見ると、22年9月期を直近期の検証であると認識した上で、直近期で精度に課題があるものの2期連続NGではないことから、継続利用が可能と結論付ける出力となった。規程を踏まえた検証リポートらしい評価を得ることができたといえる。

## 自社情報に沿った回答を得るための工夫も

今回の試行実験では、検証リポートの作成に必要なすべての情報をプロンプトに入力した。しかし、現実の規程は膨大である。プロンプトの入力の長さ制限があることや（入力にかかる）利用コストの増加、関係性の低い文章を含めることによる回答精度の低下等の理由から、すべてをプロンプトに入力することは非現実的である。また、自社情報をもとに生成AIをファインチューニングする手法も存在するが、コストや手間がかかることが多い。

そこで注目されているのが、RAG（Retrieval Augmented Generation＝検索拡張生成）である。

この技術では、まず自社情報をもとにベクトルデータベース（テキスト情報等を数値ベクトルに変換し、類似検索をしやすいとしたもの）を作成する。次に、そのベクトルデータベースからモデル検証結果と関連する規程を検索し、その検索結果をリポート自動作成時の参考情報として生成AIへ与える。これにより、ファインチューニングなしに、自社情報に沿った回答を得ることが可能になるほか、ハルシネーション（後述）の抑制にもつながる。

また、最近ではテキスト以外の情報も処理できるマルチモーダル機能が強化されつつある。規程や過去の検証リポートに含まれる図表などを認識した上で、リポートを自動作成することが可能になる日も近いだろう。

## 考え得る活用場面と検討すべき論点

試行実験の結果を踏まえると、

シンプルな検証リポートであれば、工夫次第で一定レベルのリポートの自動作成は可能といえる。また、管理すべきモデル数が多くなるほど、自動化のメリットが大きくなる。

例えば、大手金融機関においては、リスクが高くないモデル（概してそのようなモデルの数は膨大である）の検証に生成AIを活用し、業務効率化と負担軽減を図ることが考えられる。それによって、より注力すべきリスクの高いモデルの検証や当局対応等にリソースを割くことが可能になる。検証リポートのすべてに今すぐ活用できなくても、一部から活用を始めるやり方は検討に値する。

また地域金融機関においては、モデル検証のリソース制約が特に大きい。そのため、あるモデルの検証の「簡易的な勘所」を生成AIに洗い出させ、その後、本格的に人が検証を実施する余地もありそうだ（リソース制約を考慮すると、大手金融機関でもそうした活用方法は考えられる）。

ほかにも、生成AIは検証プ

ログラムの自動作成やタスクの網羅的なリストアップが得意である。従って、モニタリングや検証ツールの内製化に加えて、金融庁原則等を踏まえた態勢整備の課題の洗い出しにも活用できるのではないだろうか。

他方で、いくつか検討すべき論点も存在する。

第一に、生成AIの回答は常に正しいとは限らない（ハルシネーション）ため、実務に精通した担当者や専門家によるレビューが一定程度は必要である。もっとも、人が関与すべき範囲については、モデルのリスクやモデルタイプなどによって、さまざまな見方が存在するだろう。

第二に、生成AIを用いて検証リポートを作成できたとして、その生成AIは管理しなくてもいいのか、どのように管理すべきなのかという論点もある。仮に重要なモデルの検証リポートを自動作成した場合、リポートの内容に誤りがあったとき、悪影響は大きくなる可能性がある。第三に、リポートの自動作成の精度を上げるには「質の高い」検証リポートの蓄積がキー

になるが、本邦ではグループ／グローバルベースでこうしたリポートがまだ多くない。本邦金融機関のMRMレベルの底上げを図り、検証リポートなどのモデルに関するさまざまな情報の蓄積が大切になる。

\*\*\*

カナダのMRMガイダンスの改訂案(注4)では、モデルの定義について「インプットデータから結果を算出するプロセス(processes input data to generate results)はモデルである」とも解釈可能な提案がなされている。これは、米国のガイダンス等におけるモデルの定義と比較すると、相当広い定義である。こうした動きから、管理すべきモデル数が急激に増え、2線のリソースがますます逼迫することが想像される。2線のリソースが不足する本邦金融機関では、MRMの業務効率化は喫緊の課題である。その対応策の一つとして、本稿で提案した生成AIを活用した検証リポートの自動作成は、可能性のある取り組みといえよう。

最後に、本稿で紹介した生成AIによる検証リポートの自動作成は「MRMの土台をある程度、整備した後に活用を模索すべき」という点は付言しておきたい。土台がぐらつく中で、いたずらに業務効率化や負担軽減を図っても、うまくいくことはないだろう。

(注)1 FRB/OCC, “SUPERVISORY GUIDANCE ON MODEL RISK MANAGEMENT” (2011年)

<https://www.federalreserve.gov/supervisionreg/srletters/sr1107a1.pdf>

2 [https://www.fsa.gov.jp/news/r-3/ginkou/20211112/pdf\\_02.pdf](https://www.fsa.gov.jp/news/r-3/ginkou/20211112/pdf_02.pdf)

3 BOE/PRA, “Model risk management principles for banks” (2017年)

<https://www.bankofengland.co.uk/prudential-regulation/publication/2023/may/model-risk-management-principles-for-banks>

4 OSFI, “Draft Guideline E-23 – Model Risk Management”

<https://www.osfi-bsif.gc.ca/en/guidance/guidance-library/draft-guideline-e-23-model-risk-management>

たけかわ まさひろ

電機メーカー等を経て19年から現職。リスク・アナリティクスチームで、主に信用リスク管理領域を中心にAI・機械学習等の最先端の技術を踏まえたアドバイザリー業務を提供。KPMG Japan Trusted AIメンバー。

たなか やすひろ

日本銀行等を経て17年から現職。21年に金融庁監督局に出向し「モデル・リスク管理に関する原則」を策定。リスク・アナリティクスチームで、主にモデルリスク管理やRRP(再建・破綻処理計画)に関するアドバイザリー業務を提供。KPMG Japan Trusted AIメンバー。

そがへ あつし

大手邦銀等を経て03年から現職。信用や市場リスク、時価評価、ストレステスト、予想信用損失会計、RAF等の領域をカバーするリスク・アナリティクスチームおよび気候変動・バーゼル規制を含めたFRMグループを統括。