



機械学習の保険実務への活用 ～過去・現在・未来～

Willis Towers Watson

- Neil Chapman (ニール・チャップマン)
- 山田 剛

日本アクチュアリー会 2022年度 第8回例会

2023年2月28日

wtw.com

© 2023 WTW. Proprietary and confidential. For WTW and WTW client use only.

wtw

機械学習の保険実務への活用 ～過去・現在・未来～

1. 過去 – 保険業界にて機械学習（ML: Machine Learning）がどのように進化してきたか？
2. 現在 – 保険業界にて、機械学習がどのように利用されているか？
3. 未来 – 今後どのように機械学習が保険業界に貢献しうるか？

過去 -
保険業界にて機械学習が
どのように進化してきたか？

機械学習とは?

専門家に尋ねよう…

“Machine learning is a subfield of artificial intelligence that allows computer systems to automatically improve their performance based on experience, without being explicitly programmed.

It involves learning algorithms that can extract insights and make predictions or decisions from data inputs.”

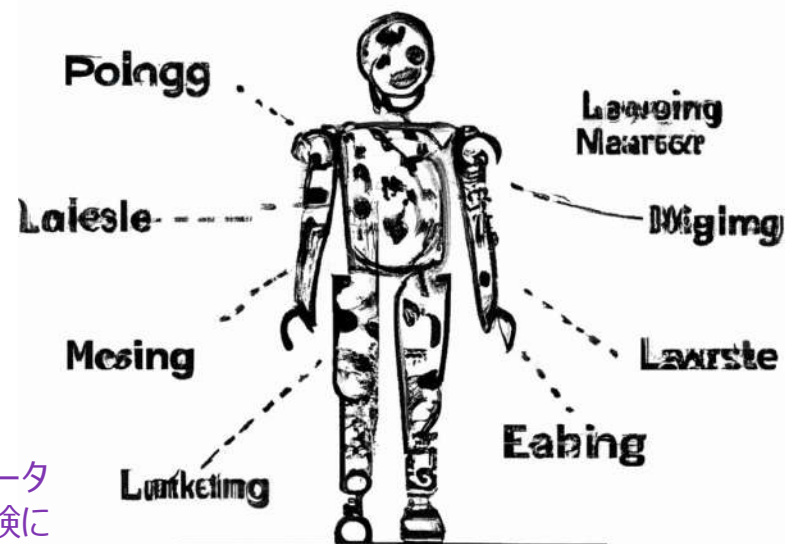


DeepL 翻訳ツール

機械学習は人工知能の一分野であり、コンピュータシステムが明示的にプログラムされることなく、経験に基づいて自動的にその性能を向上させることを可能にする。

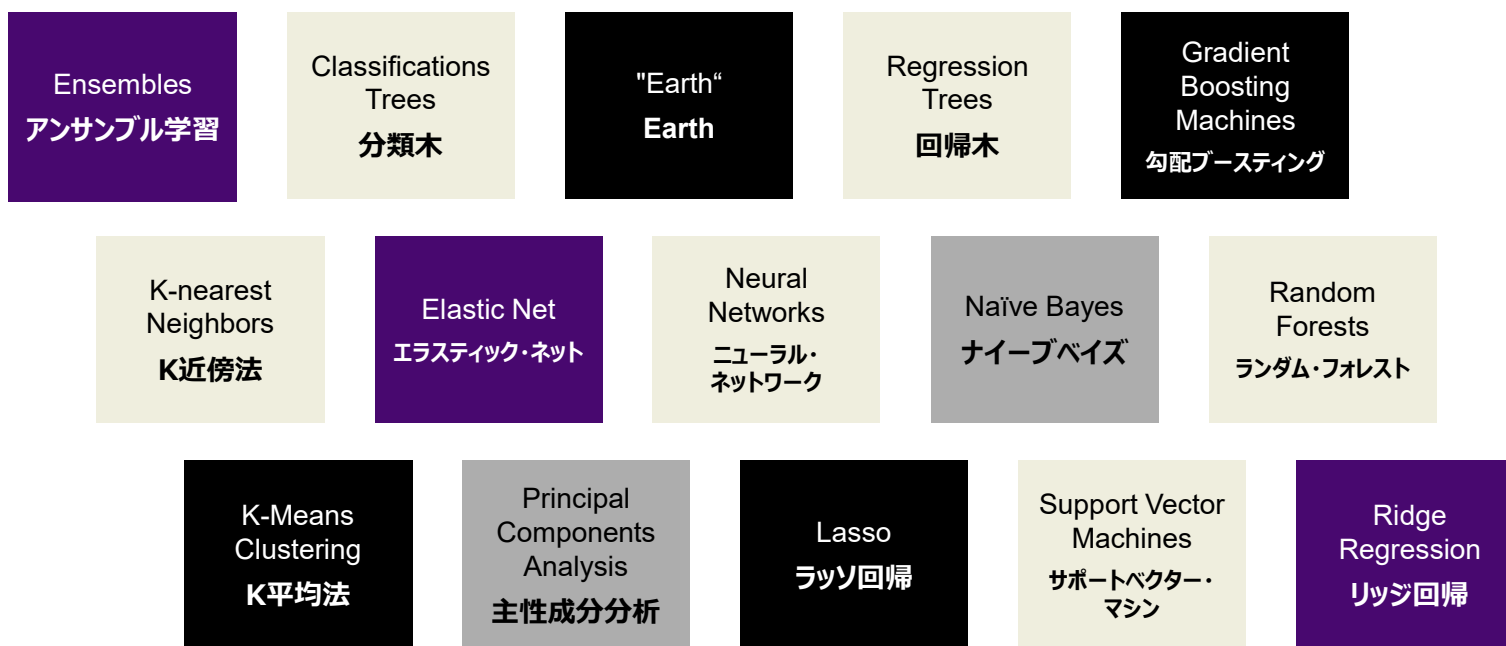
機械学習では、データ入力から洞察を引き出し、予測や決定を行うことができるアルゴリズムを学習する。

Learning Marketing

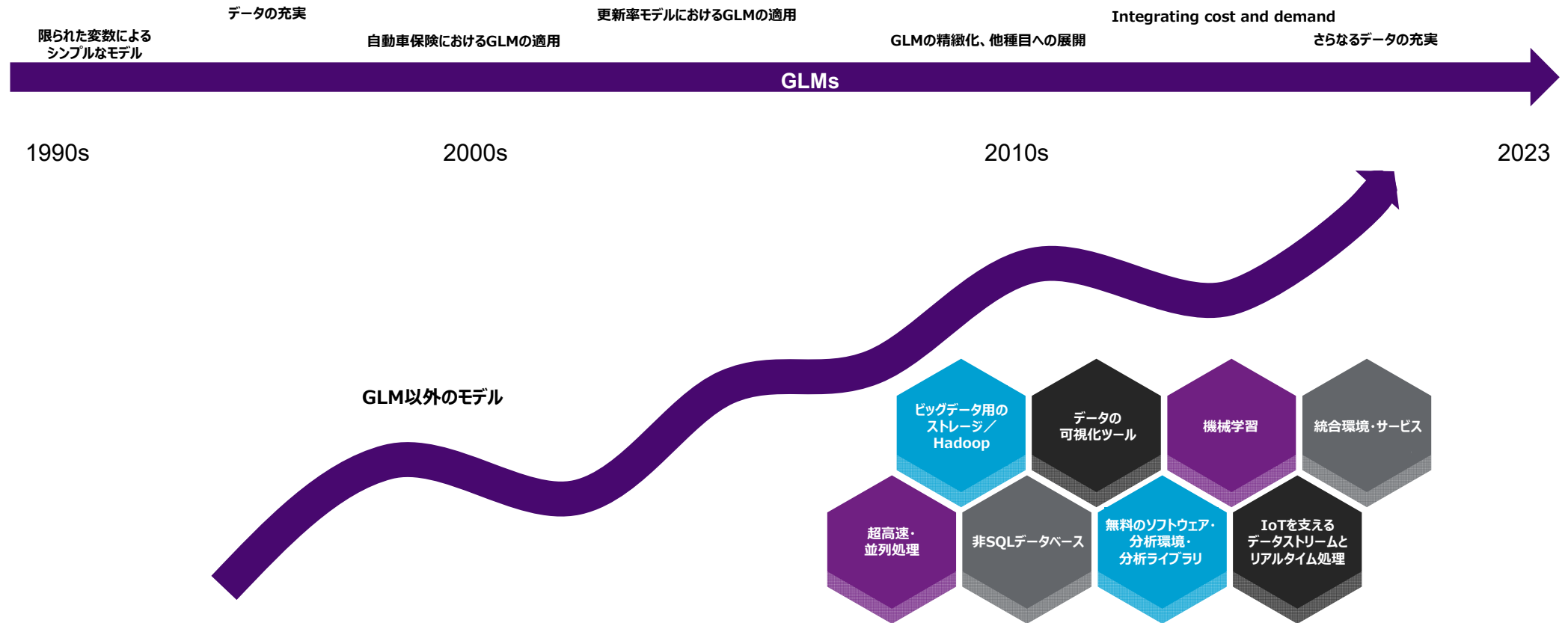


DALL-E

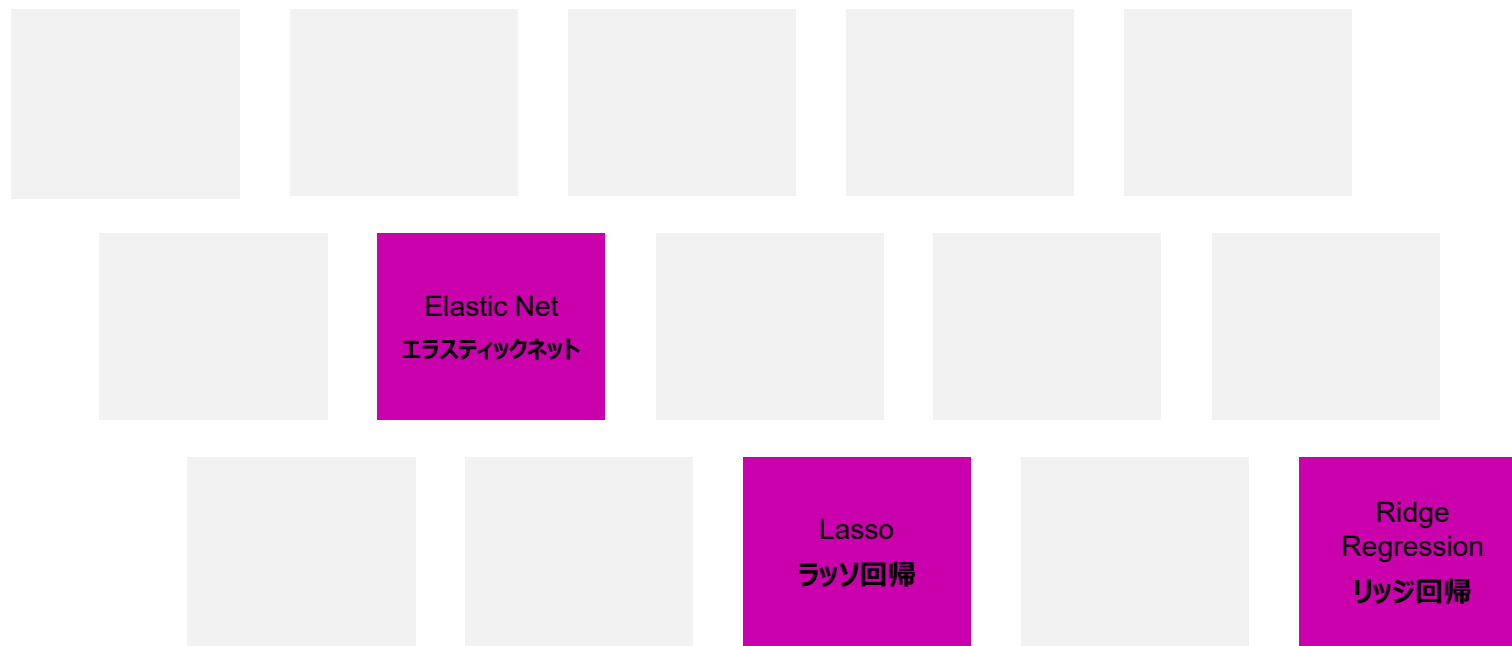
これらの機械学習手法とは?



新しいものではないが……



一例として正則化回帰モデルの解説



Penalized regression (罰則付き・正則化回帰)

概要

GLMs

- 予測値 $f(x) = g^{-1}(X \cdot \beta)$
- β は損失関数 $L(\beta|X, y)$ を最小化するように算出される (X 説明変数の観測値, y 目的変数の観測値)

Penalized regression (罰則付き・正則化回帰)

- 損失関数…… $L(\beta|X, y) + \lambda$ “損失関数に β に関する罰則項を付加”

Elastic Net

$$\text{最小化: } L(\beta|X, y) + \lambda_1 \sum_i |\beta_i| + \lambda_2 \sum_i \beta_i^2$$

Lasso回帰 – 青四角部分 * Lasso : Least Absolute Shrinkage Selection Operator

- 重要でないパラメタをゼロにするペナルティ (罰則) 項であり、変数選択に有効

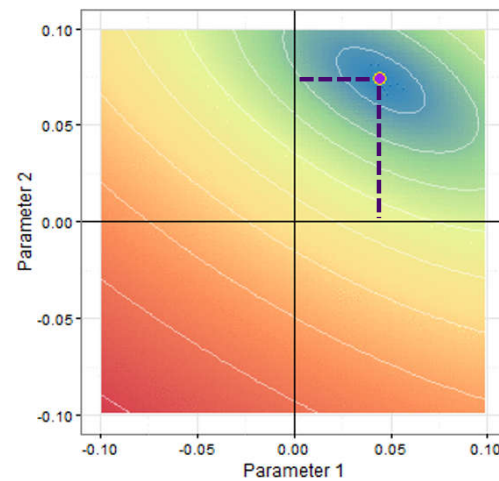
Ridge回帰 – 紫四角部分

- 極端なパラメタにペナルティを加えるものの、パラメタをゼロにするものではない

Penalized Regression (罰則付き・正則化回帰)

GLM

$f(\underline{x}) = g^{-1}(\mathbf{X}\cdot\beta)$ 但し β は右式を最小化するように決定 $L(\beta|X, y)$



Penalized Regression (罰則付き・正則化回帰)

$f(\underline{x}) = g^{-1}(\mathbf{X} \cdot \underline{\beta})$ 但し $\underline{\beta}$ は右式を最小化するように決定

GLM

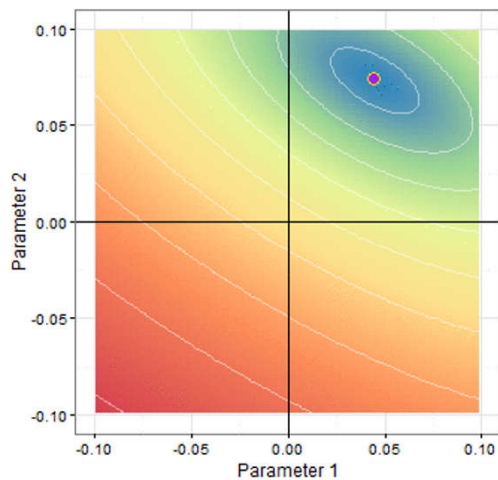
Lasso

Ridge

$$L(\beta|X, y) + \lambda_1 \sum_i |\beta_i| + \lambda_2 \sum_i \beta_i^2$$

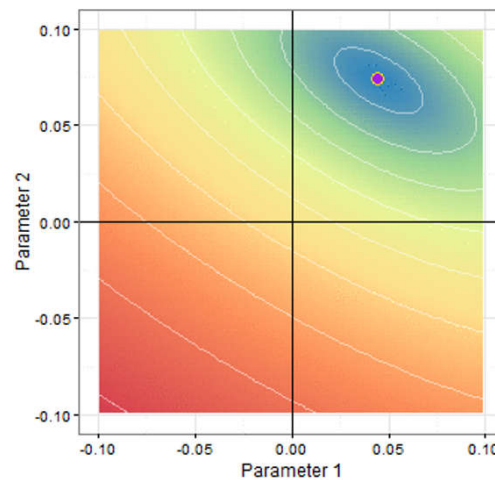
Elastic Net

Ridge $\sum_i \beta_i^2$



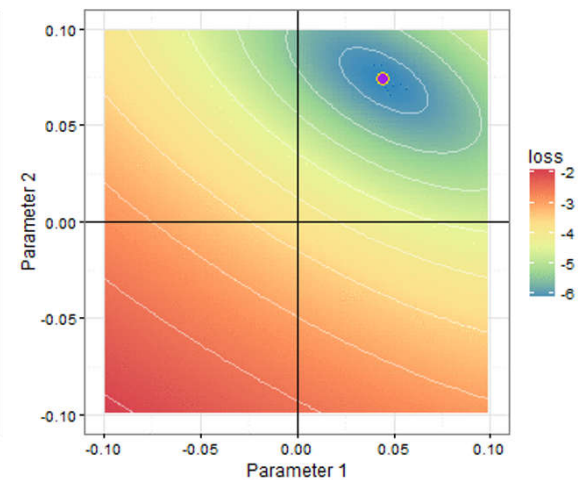
極端なパラメタにペナルティを加えるものの、
パラメタをゼロにするものではない

Elastic Net



両者の組み合わせ

Lasso $\sum_i |\beta_i|$



重要でないパラメタをゼロにするペナルティ (罰
則) 項であり、変数選択に有効

Penalized Regression (罰則付き・正則化回帰)

$f(\underline{x}) = g^{-1}(\mathbf{X} \cdot \underline{\beta})$ 但し $\underline{\beta}$ は右式を最小化するように決定

GLM

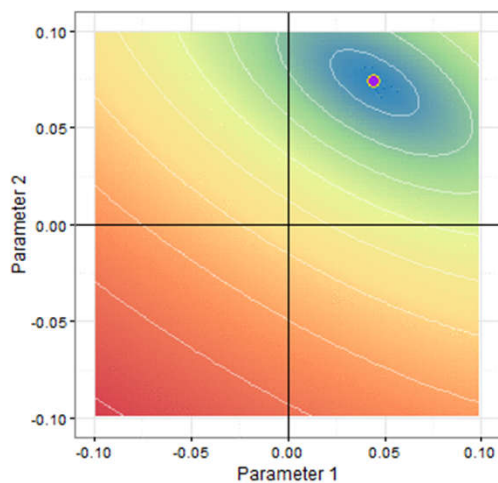
Lasso

Ridge

$$L(\beta|X, y) + \lambda_1 \sum_i |\beta_i| + \lambda_2 \sum_i \beta_i^2$$

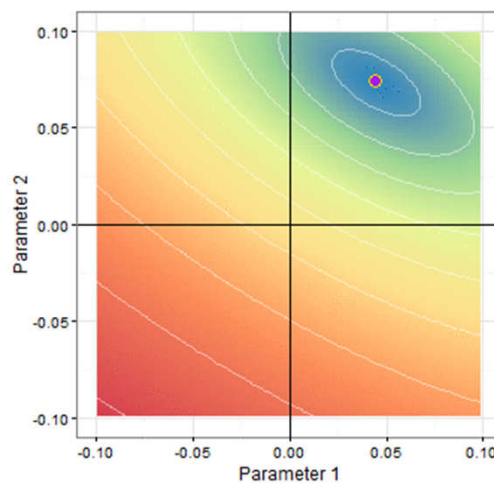
Elastic Net

Ridge $\sum_i \beta_i^2$



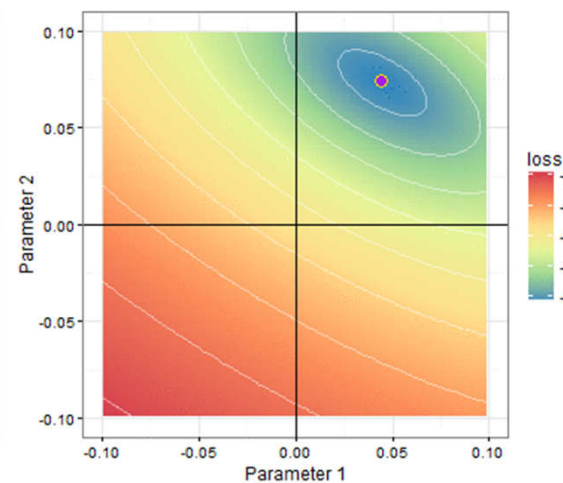
極端なパラメタにペナルティを加えるものの、
パラメタをゼロにするものではない

Elastic Net



両者の組み合わせ

Lasso $\sum_i |\beta_i|$

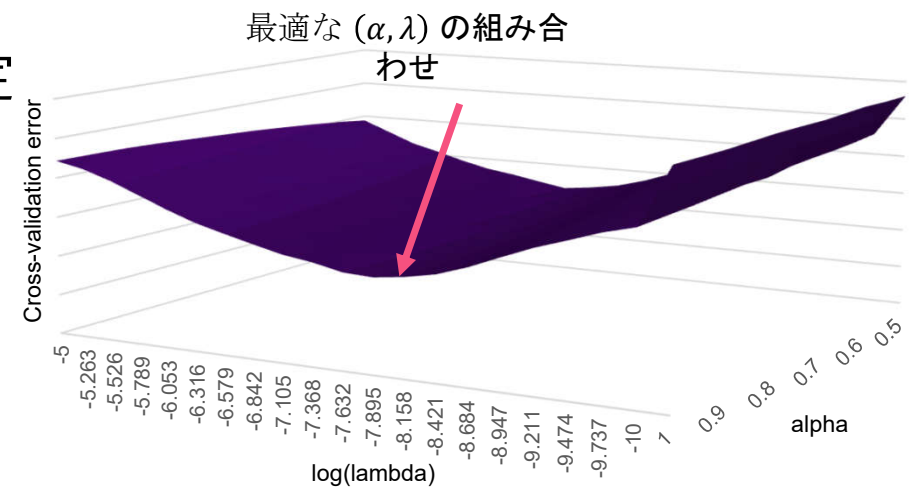


重要でないパラメタをゼロにするペナルティ (罰
則) 項であり、変数選択に有効

Penalized regression (罰則付き・正則化回帰)

パラメタ選択

- 最小化: $L(\beta|X, y) + \lambda_1 \sum_i |\beta_i| + \lambda_2 \sum_i \beta_i^2$
- 罰則パラメタは右の通り表現できる: $\lambda_1 = \lambda\alpha, \quad \lambda_2 = \lambda \left(\frac{1-\alpha}{2} \right)$
- α は Lasso ($\alpha = 1$) と Ridge ($\alpha = 0$) のバランスを司る
- λ は 全体的な罰則化の強さを司る
- λ, α は交差検証(cross-validation)により決定
- 係数は初期設定より決定される



Configuration

Model Fitting

Validation

Model Parameters

GLM

Model Term Selection

Show Banding

- Time
- Age_of_Main_Driver
- Age_of_Youngest_Driver
- Age_of_Youngest_Additional_Driver
- Gender_of_Main_Driver
- Gender_of_Youngest_Driver
- Gender_of_Youngest_Additional_Driver
- Minimum_Licence_Held_Years
- Marital_Status_Main_Driver
- Driving_Restriction
- Number_of_Drivers
- Vehicle_Age
- Vehicle_Group
- Make
- Vehicle_Value
- Use
- Annual_Mileage
- Rating_Area
- Voluntary_Excess
- Bonus_Malus_Years
- Bonus_Malus_Protection
- Number_of_Past_Claims
- Number_of_Past_Convictions
- Policy_Tenure
- Payment_Method
- Payment_Frequency
- Random10

Penalisation Parameters

Name	Value
Alpha	0
Lambda	0

Tuning

GLM

Model Registry

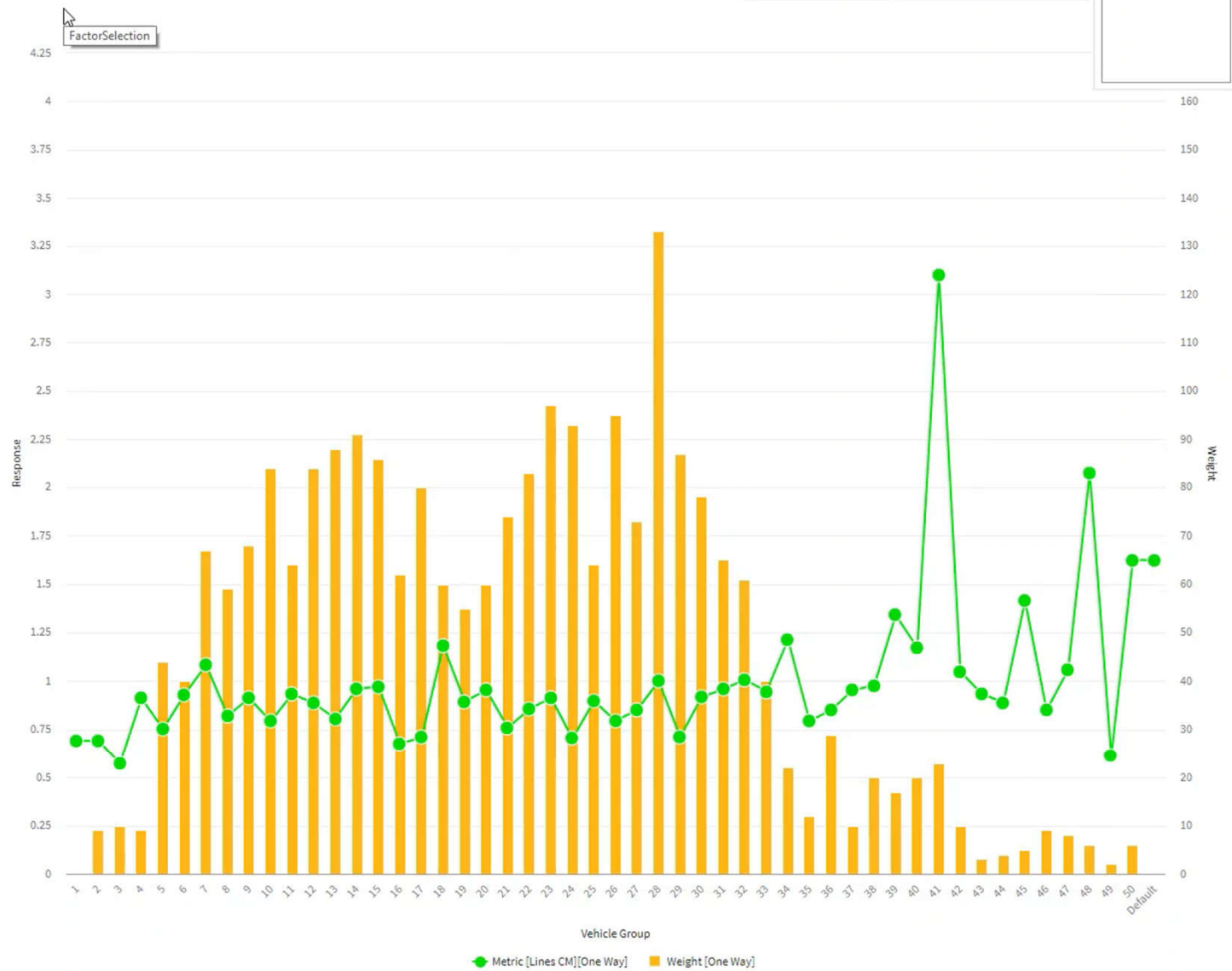
Save Fitted Model

First term: Second term:













Rescale Switch

Current Reference Switch

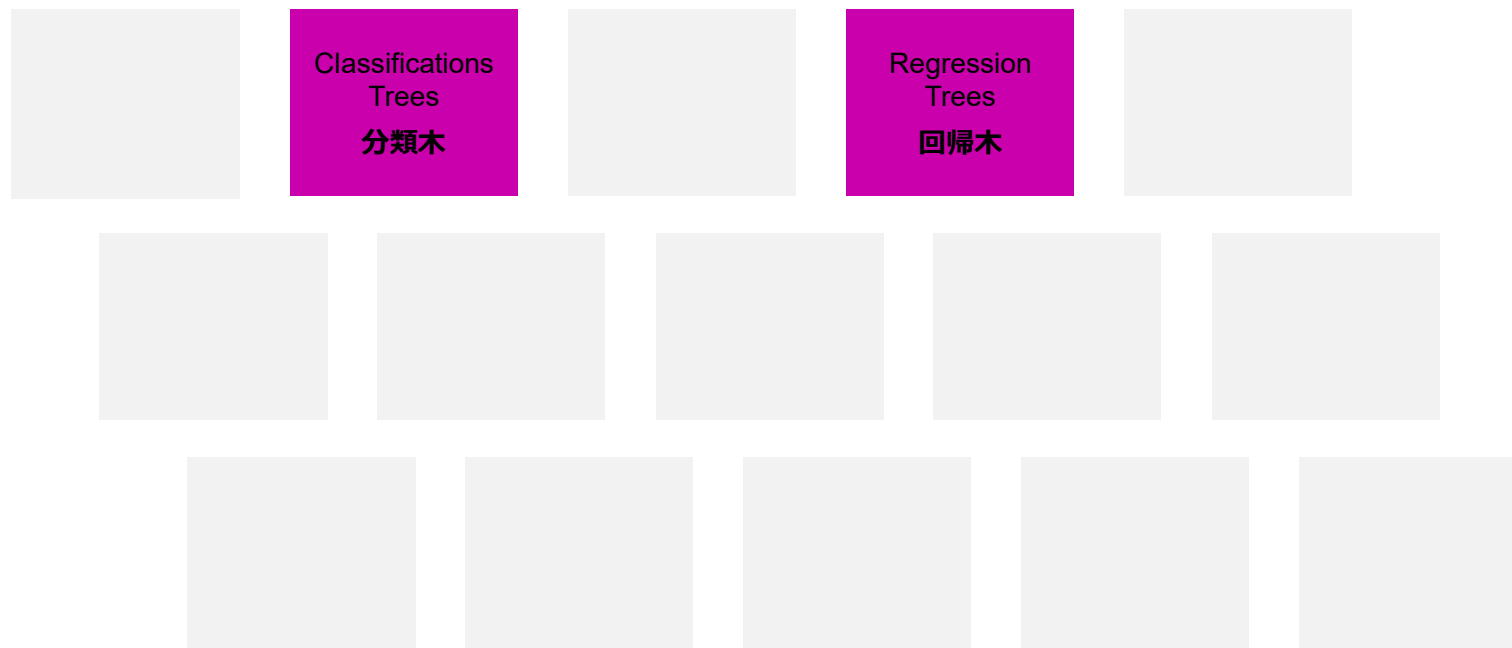
Metric Selection



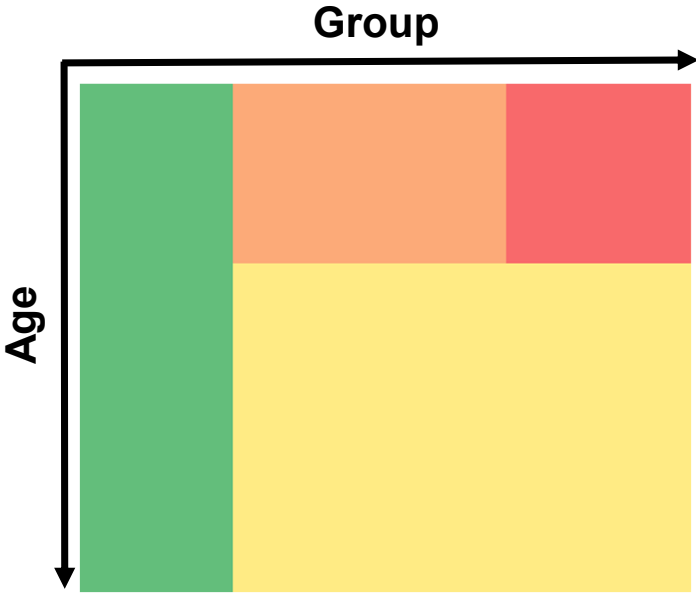
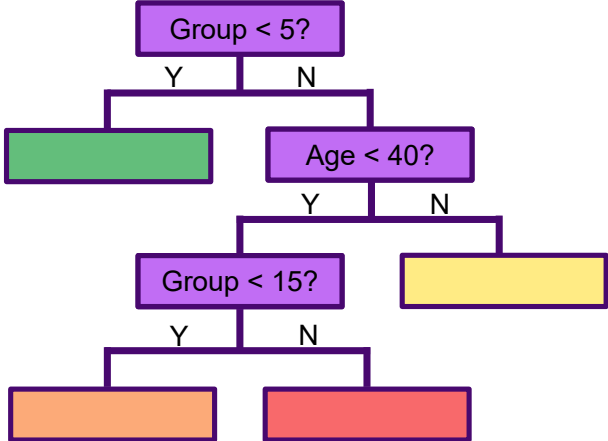
Penalized regression (罰則付き・正則化回帰)

	GLM	Penalized regression
Predictive power 予測力		
Interpretation 解釈性		
Implementation 実行		
Stability 安定性		
Execution speed 実行速度		
Analytical time and effort 分析時間・ロード		

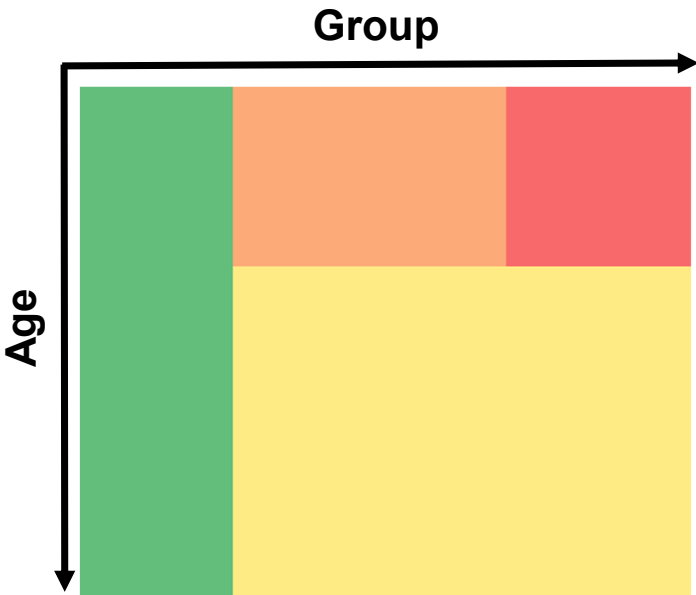
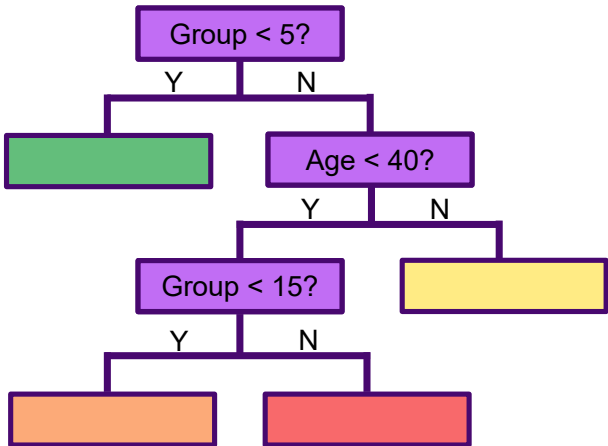
様々な“木”



Decision Trees (決定木)

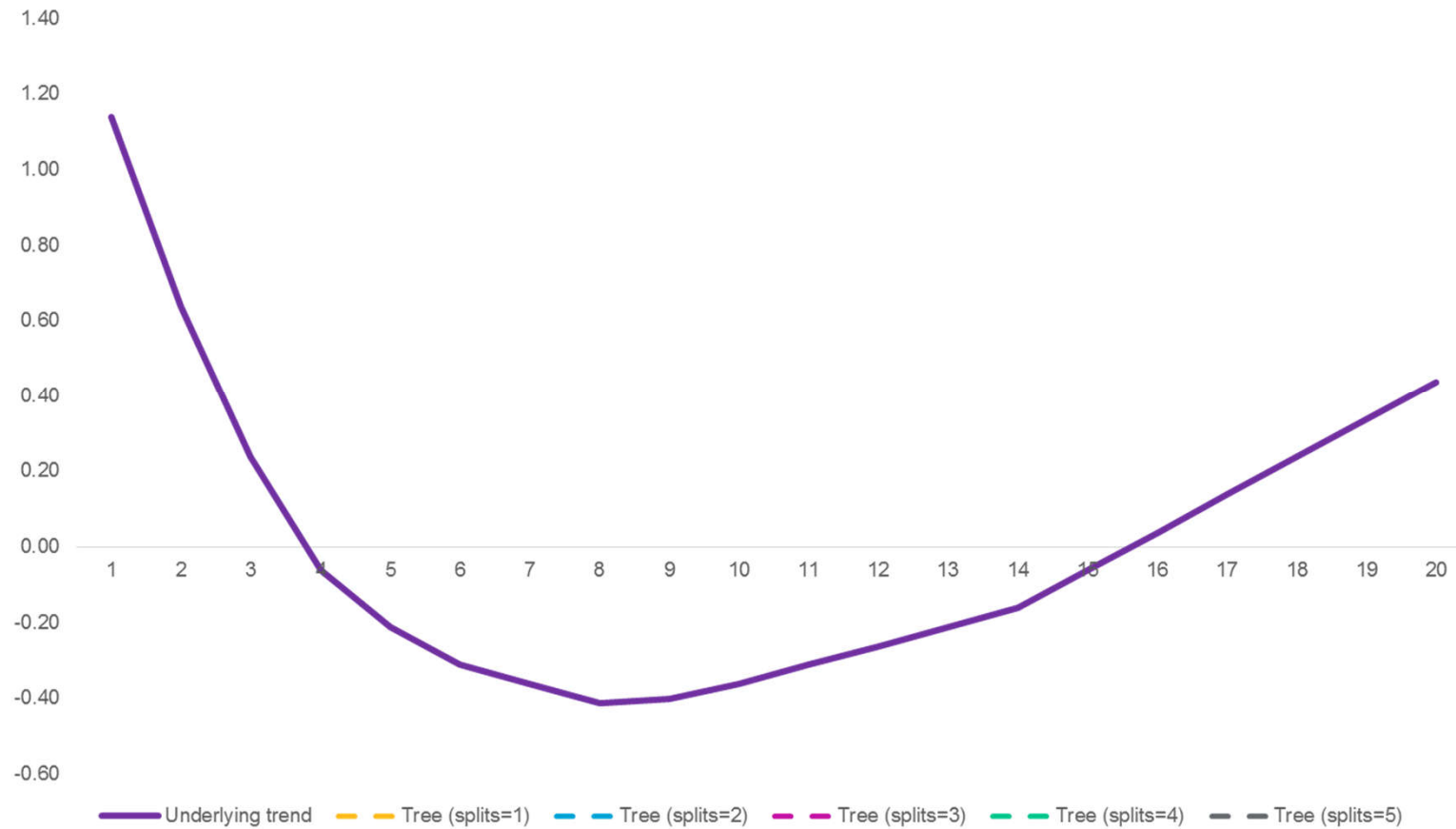


Decision Trees (決定木)



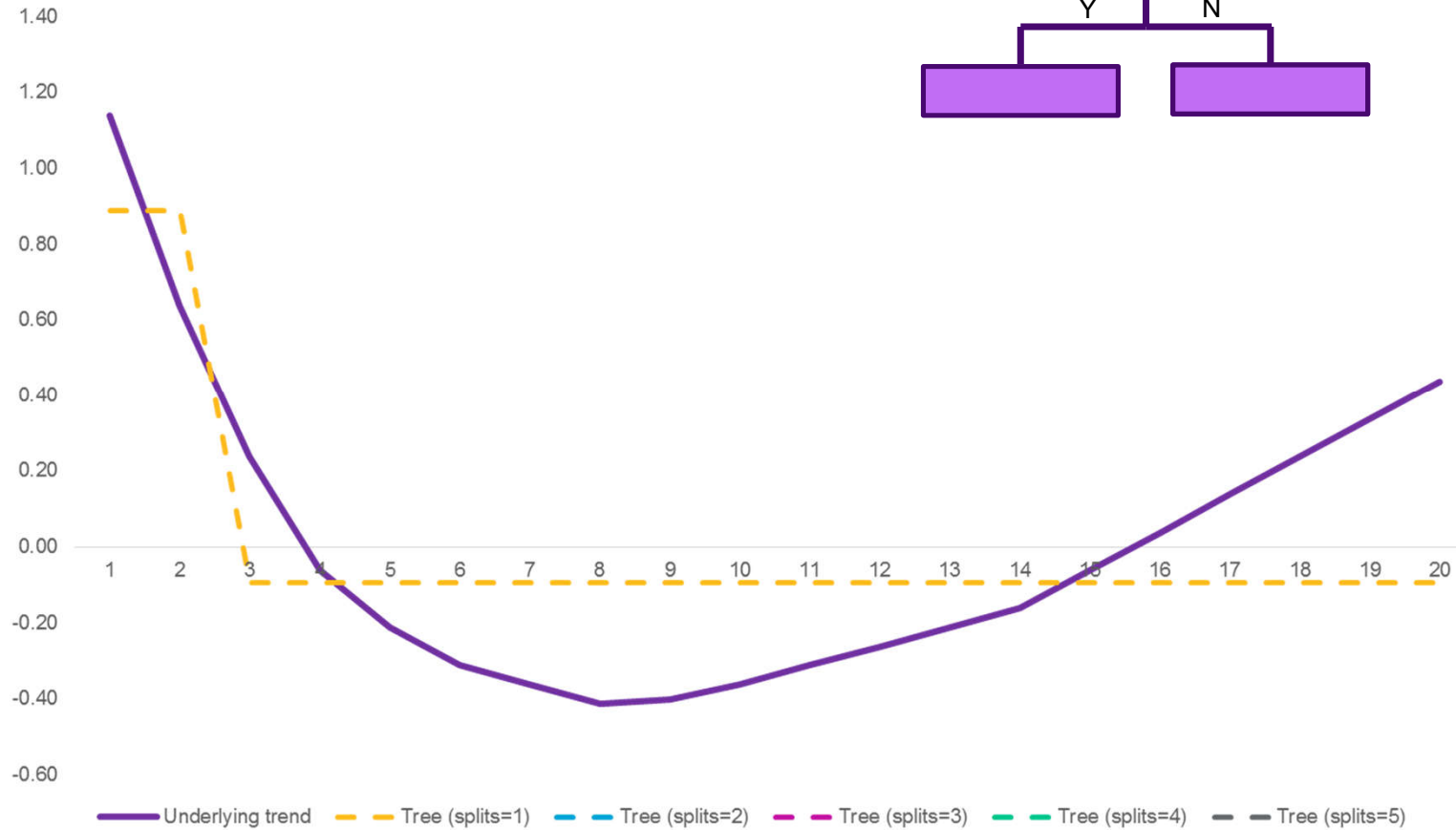
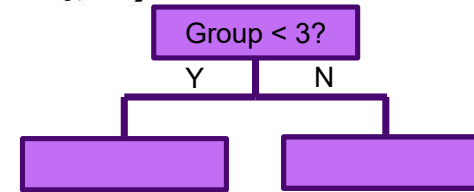
A simple Tree example (単純な木の例)

Tree results



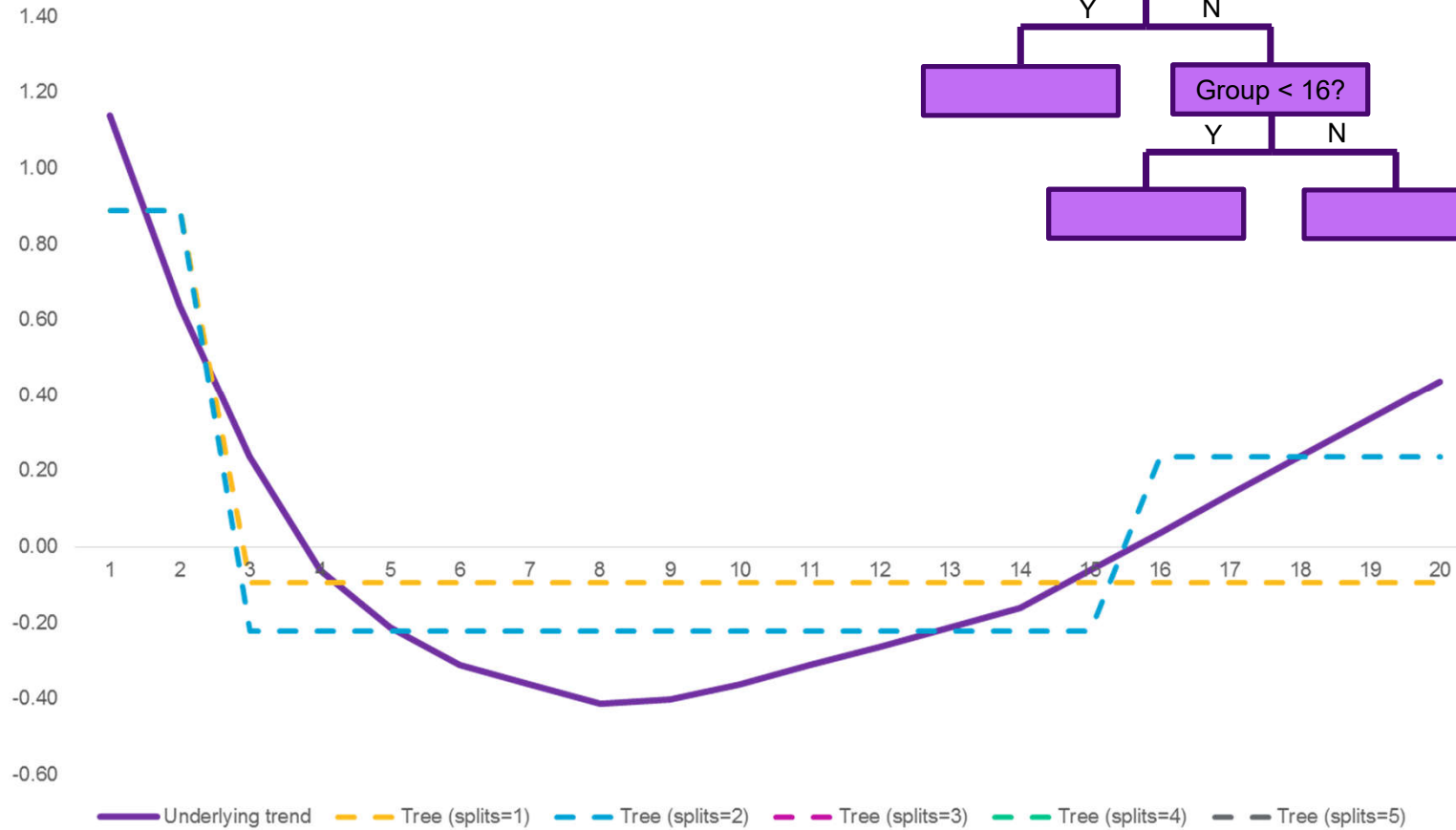
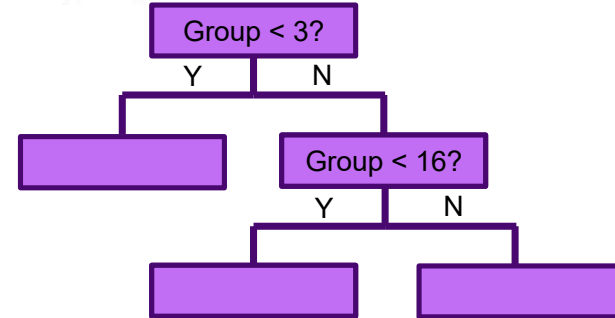
A simple Tree example (単純な木の例)

Tree results



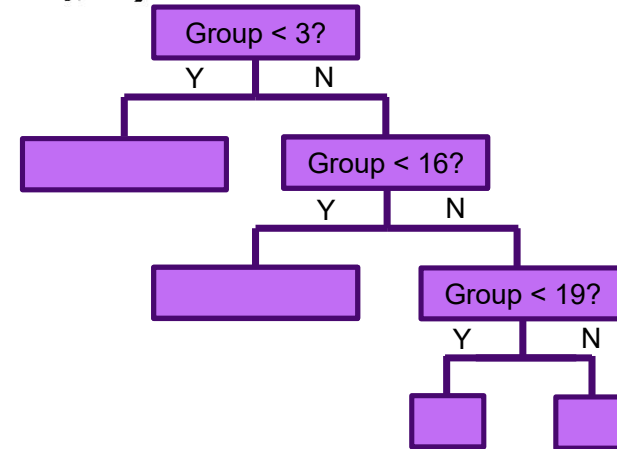
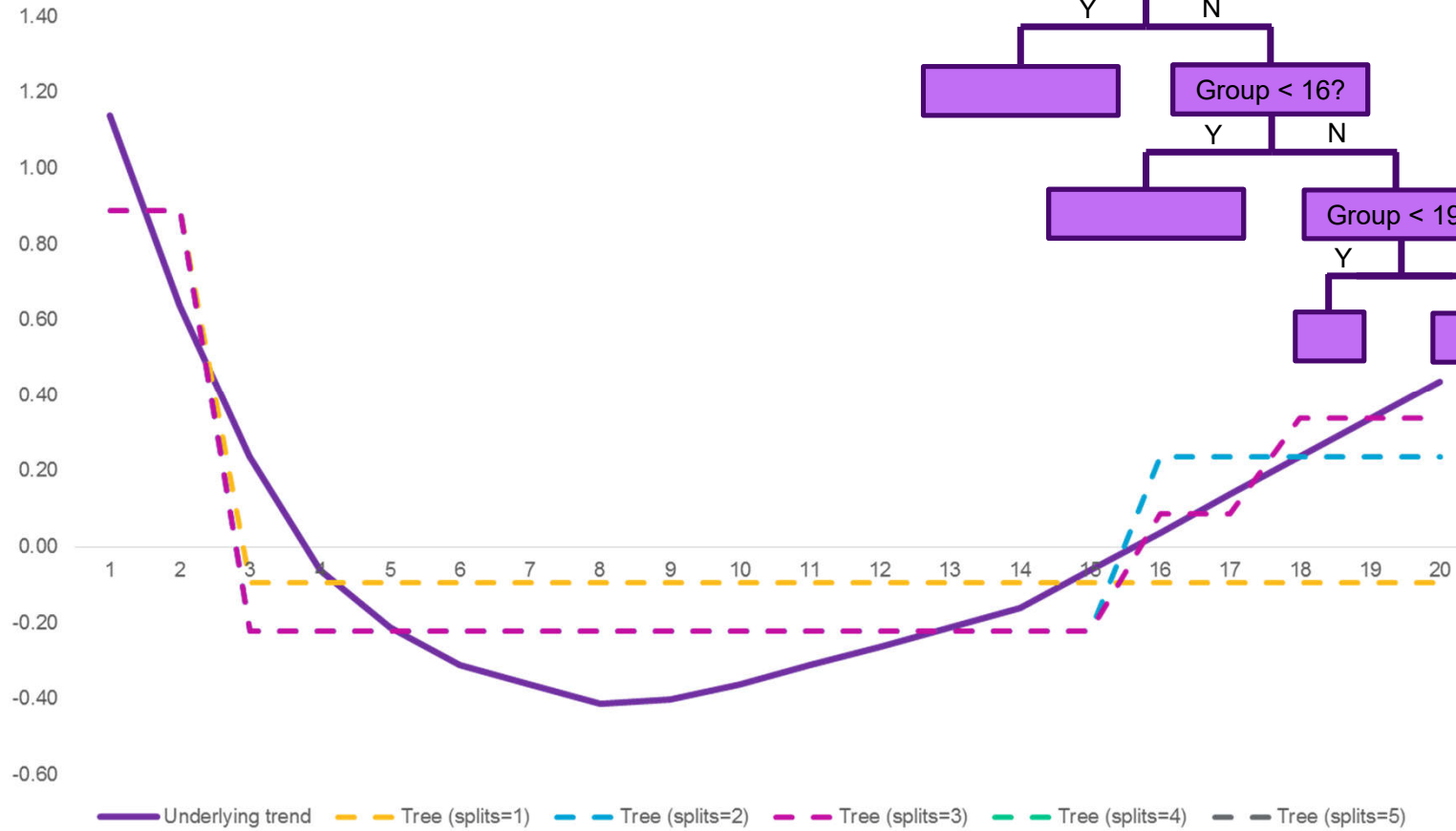
A simple Tree example (単純な木の例)

Tree results





















A simple Tree example (単純な木の例)

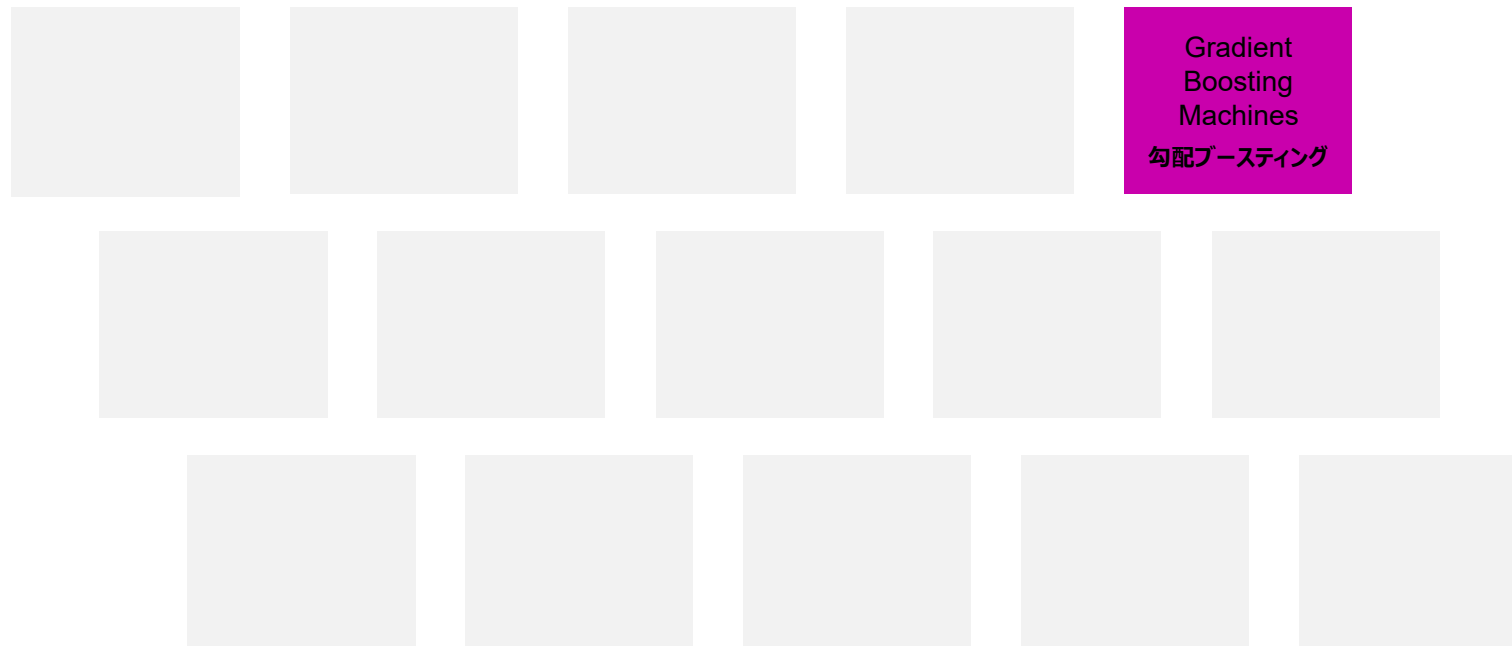
Tree results



Decision trees (決定木)

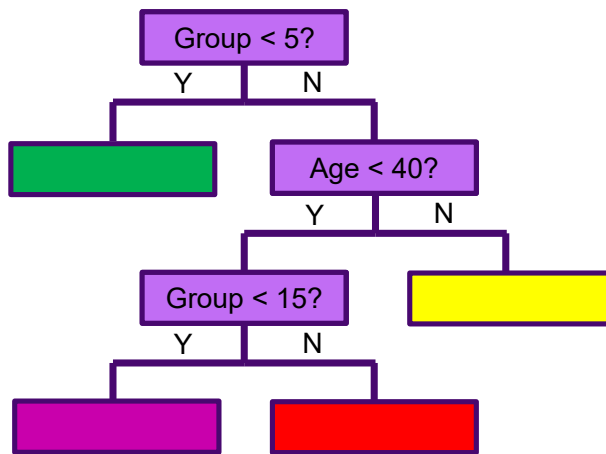
	GLM	Penalized regression	Decision trees
Predictive power 予測力			
Interpretation 解釈性			
Implementation 実行			
Stability 安定性			
Execution speed 実行速度			
Analytical time and effort 分析時間・ロード			

Gradient Boosting Machines (勾配ブースティング)



Gradient Boosted Machine or “GBM”

A tree
 $f_i(x)$



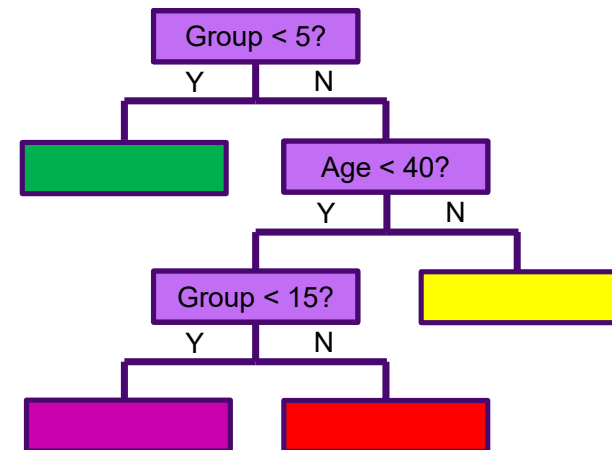
A GBM

$$f(x) = \lambda \sum_{n=1}^N f_n(x)$$



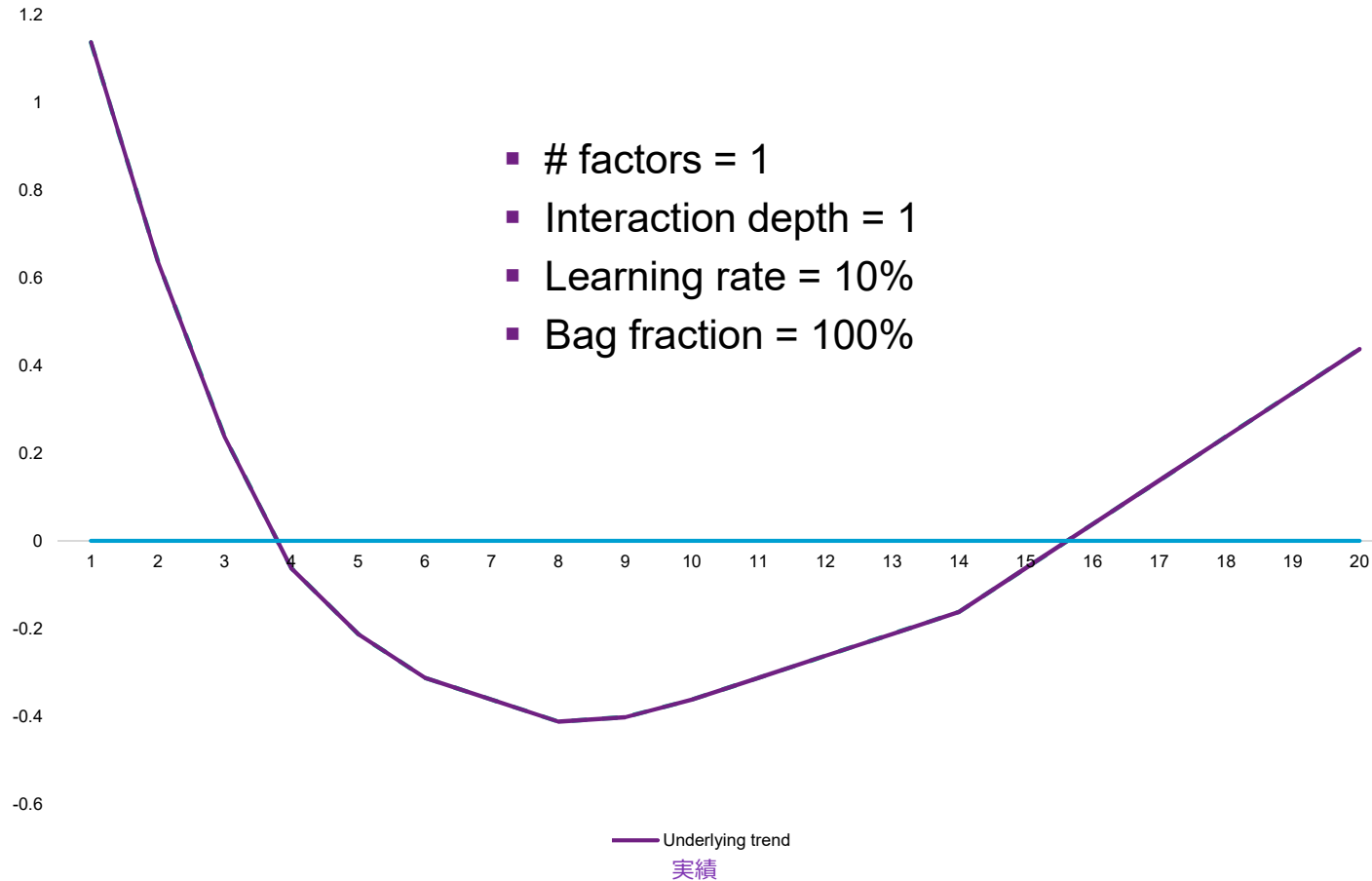
4つの設定値

- λ : **Learning Rate 学習率 / “縮小化”**
 - 更新前の予測値に対して、次の分岐計算の際にどの程度数値をアップデートするか
 - 更新後の予測値 =
更新前の予測値 + (予測値 × 学習率)
- **Interaction Depth : 相互作用深度**
 - 各木における分岐の最大深度 (または、分岐の数 - 1)
- **N : 決定木の本数 (繰り返し数)**
- **Bag fraction (サンプリング割合)**
 - Row Sampling … 各々の決定木において、一部のデータをランダムに利用することで (Bag fraction)、モデルが安定する。
 - Column Sampling … 各々の決定木において、説明変数をランダムに利用することで、モデルが安定化する。



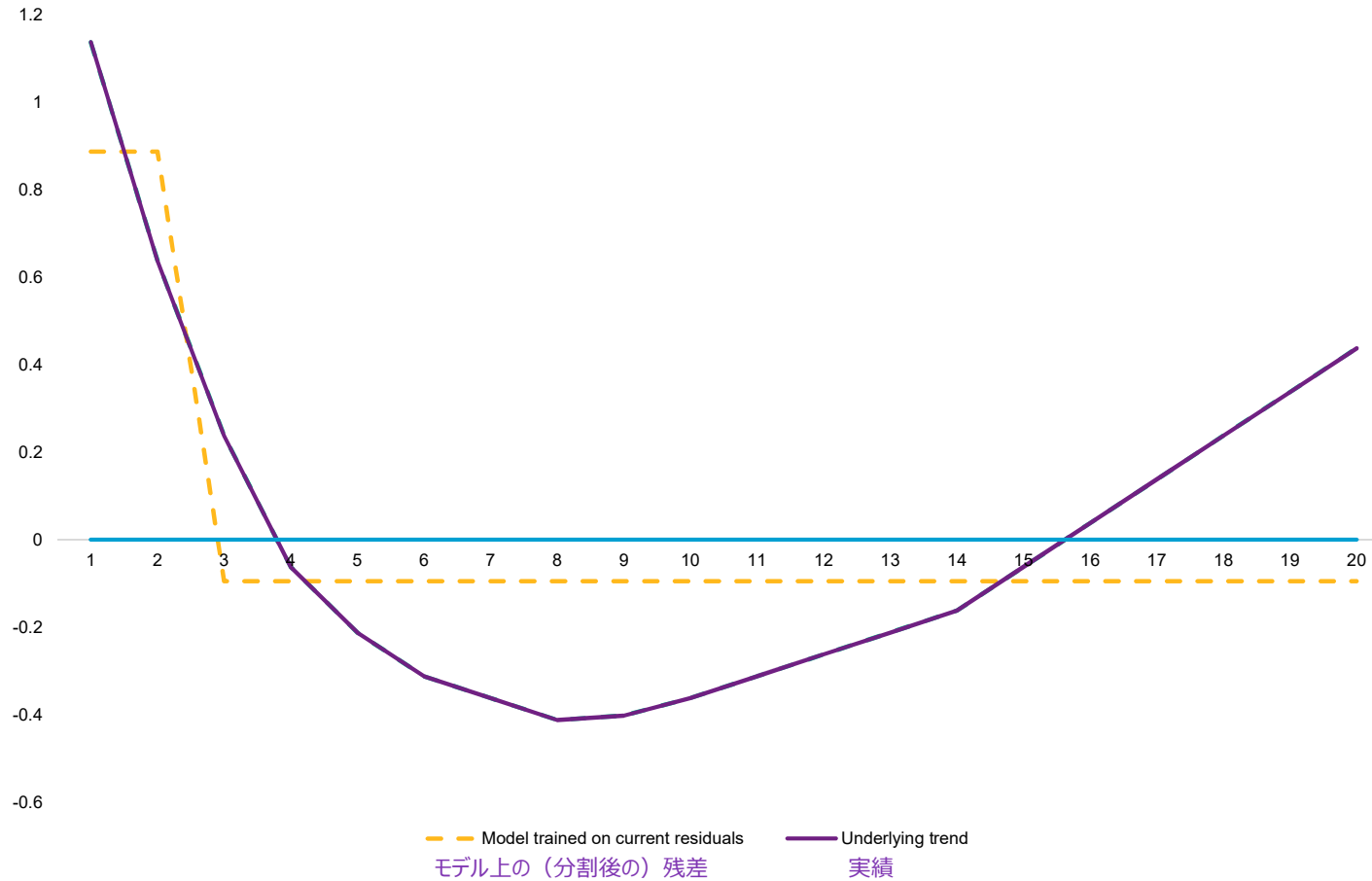
A simple GBM example (シンプルなGBMの例)

GBM results at iteration 0



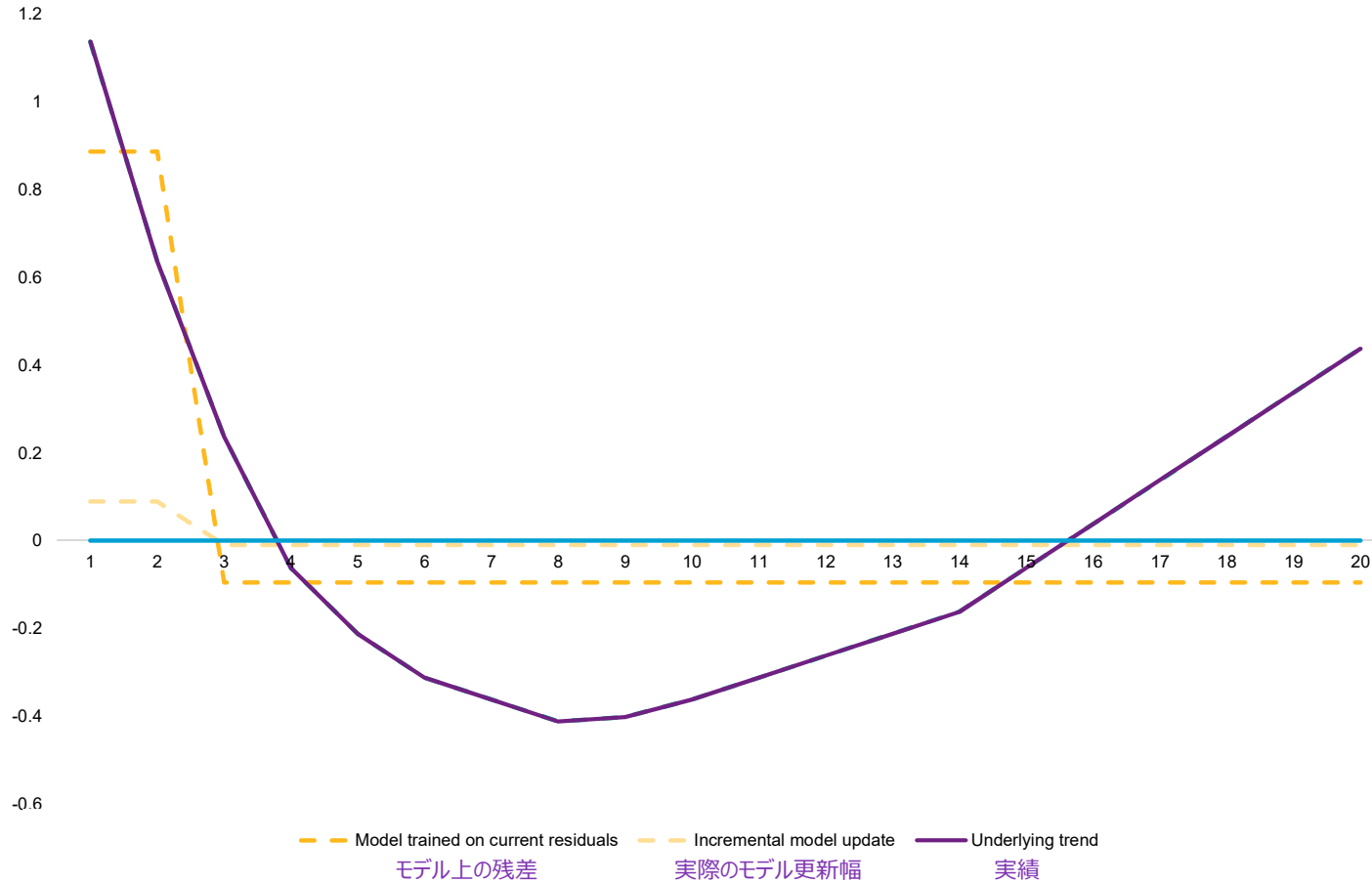
A simple GBM example (シンプルなGBMの例)

GBM results at iteration 0



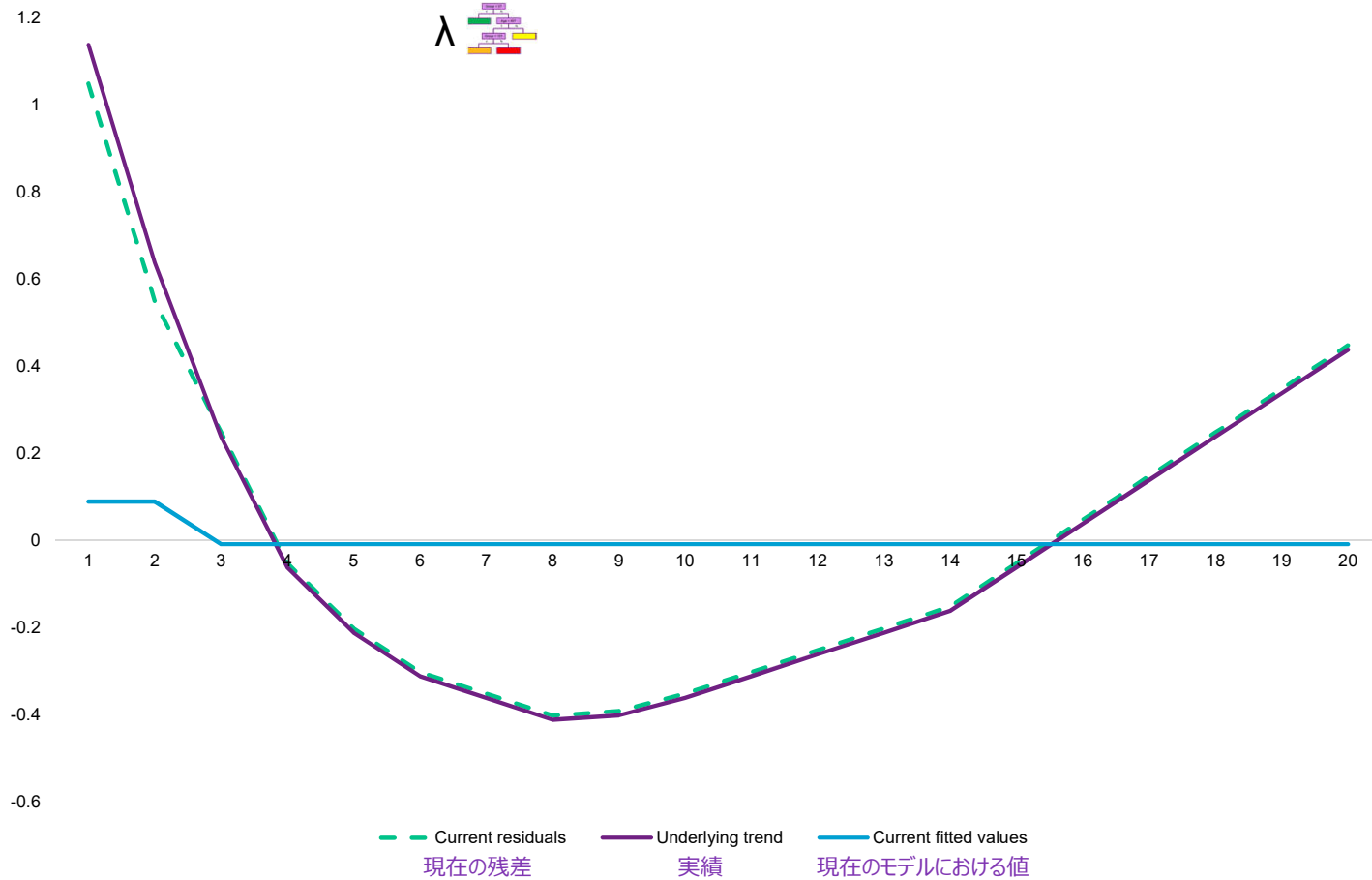
A simple GBM example (シンプルなGBMの例)

GBM results at iteration 0

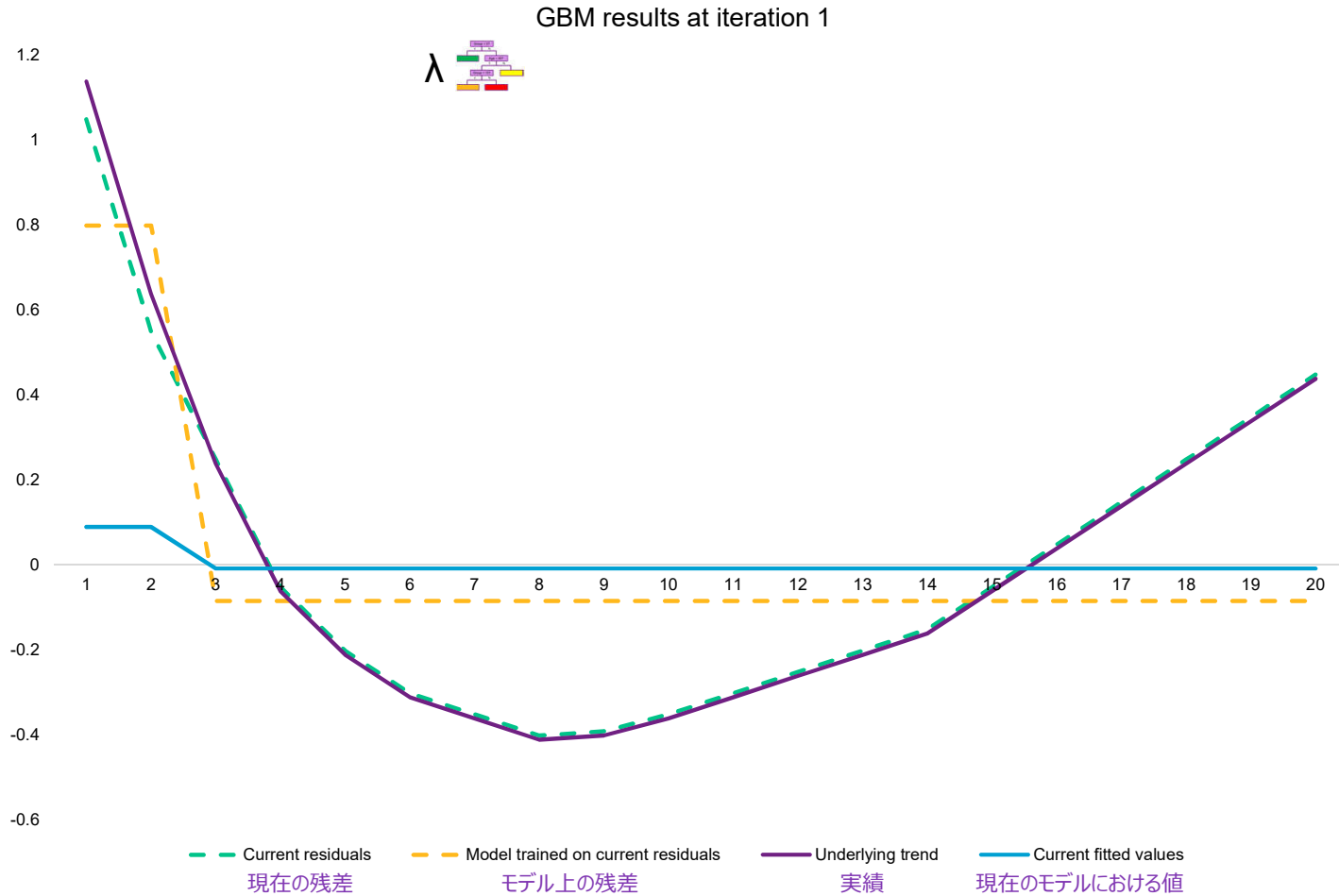


A simple GBM example (シンプルなGBMの例)

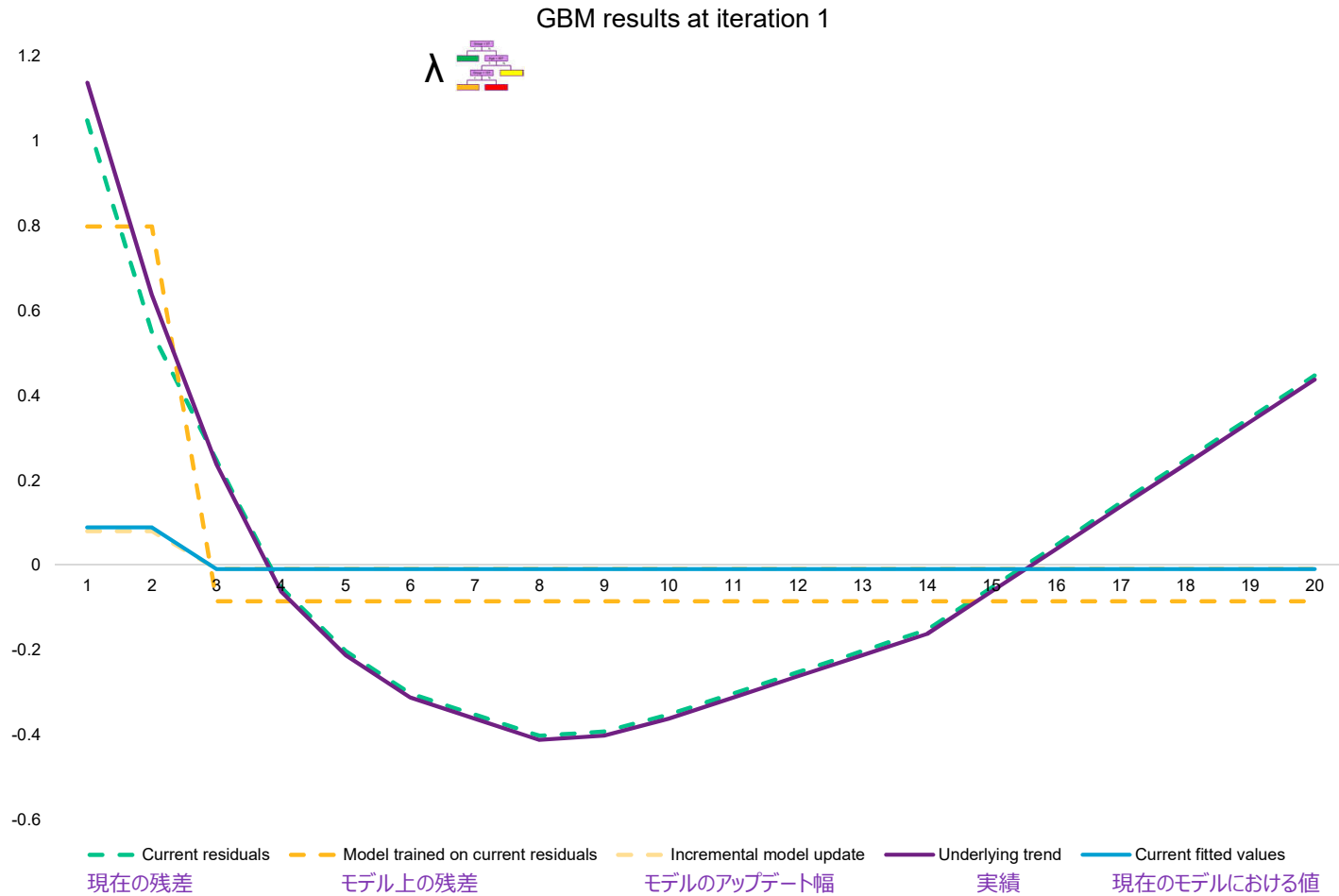
GBM results at iteration 1



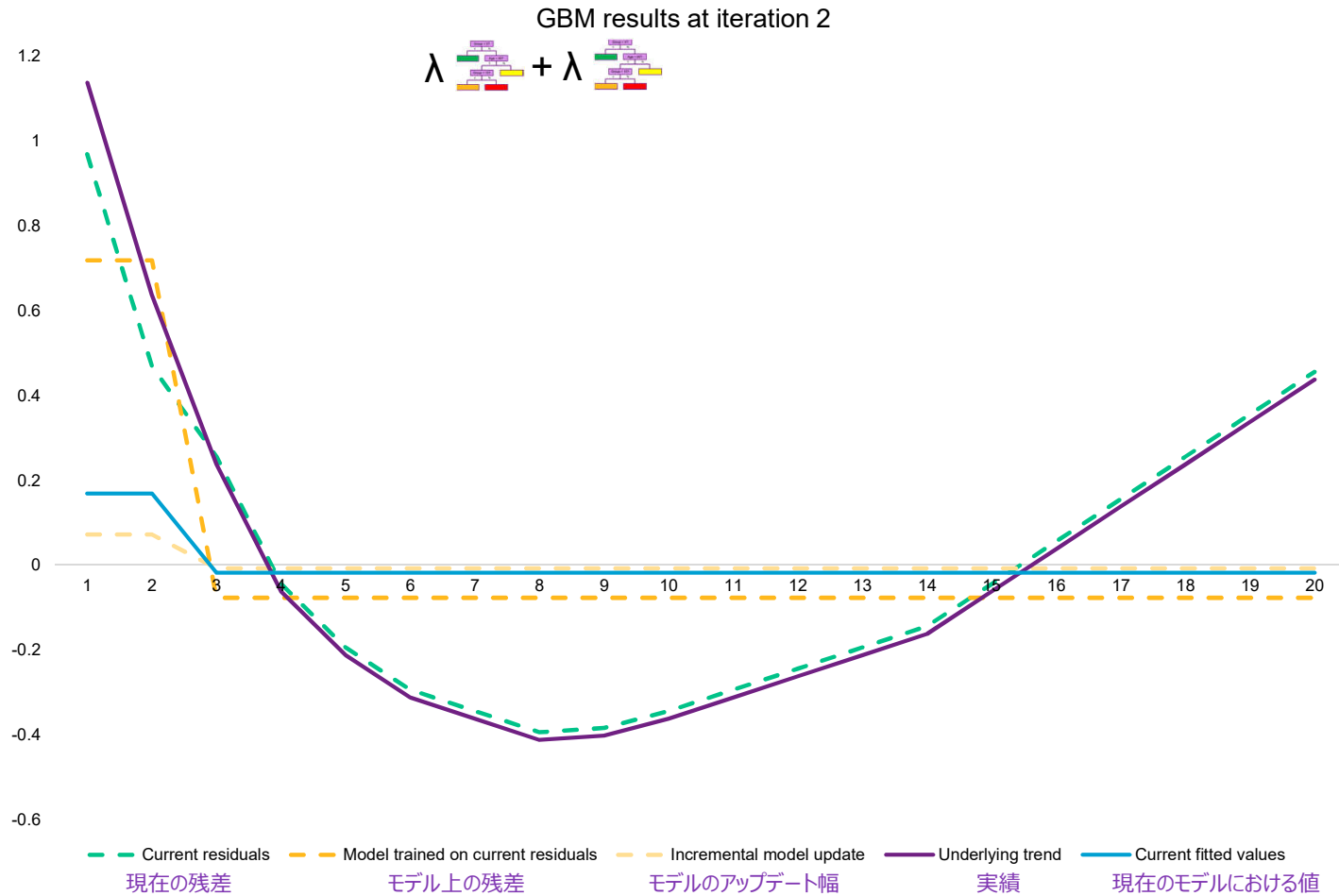
A simple GBM example (シンプルなGBMの例)



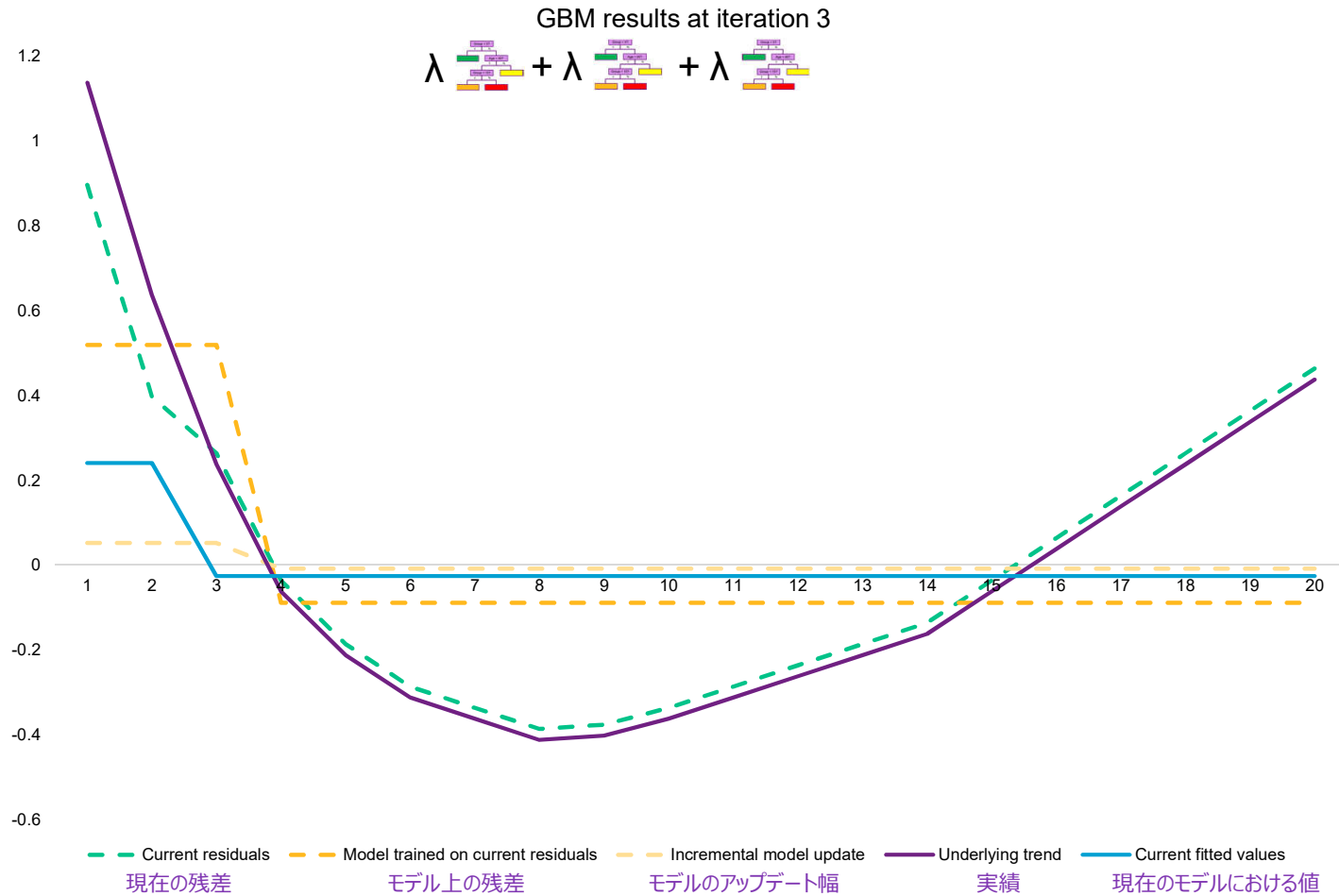
A simple GBM example (シンプルなGBMの例)



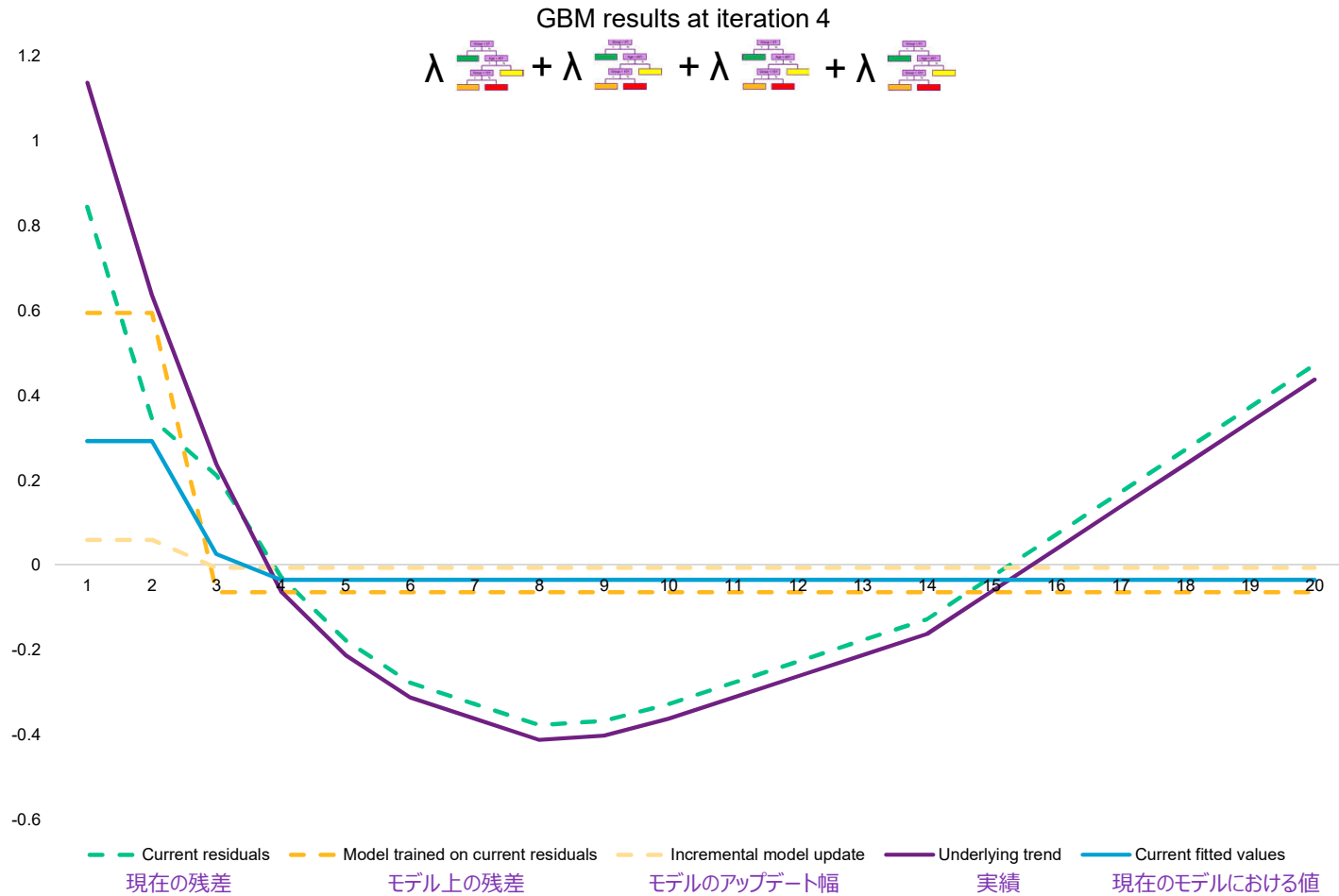
A simple GBM example (シンプルなGBMの例)



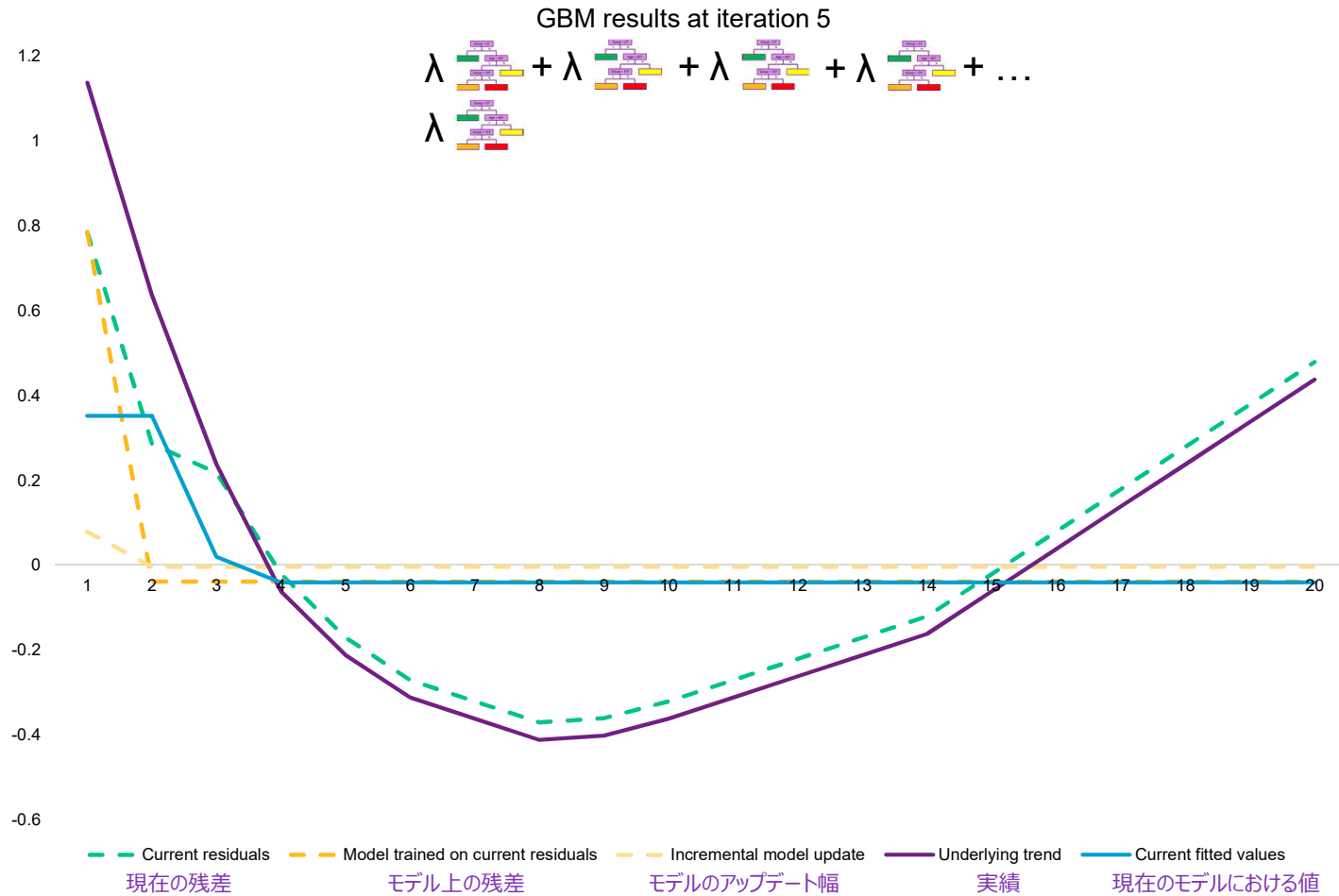
A simple GBM example (シンプルなGBMの例)



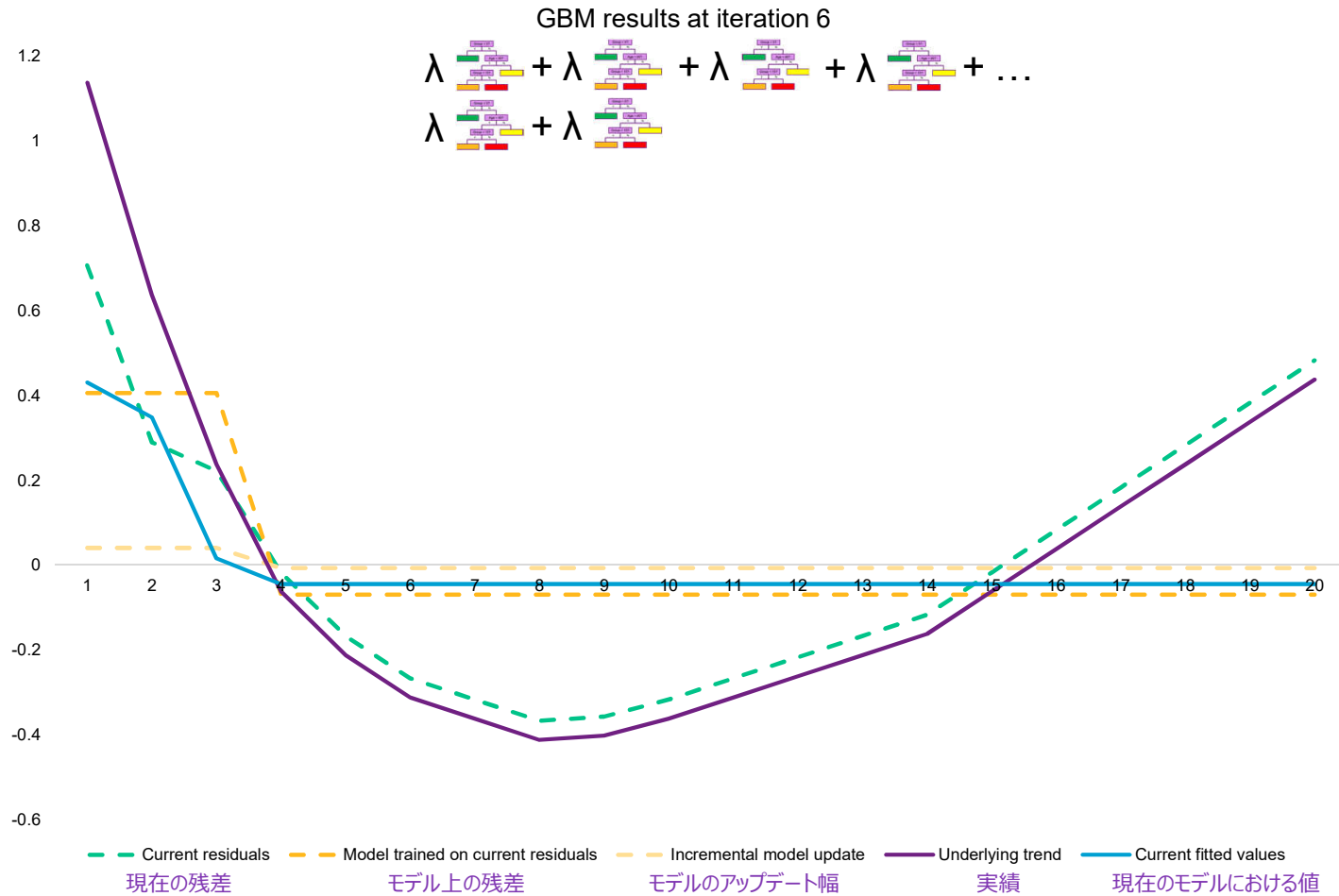
A simple GBM example (シンプルなGBMの例)



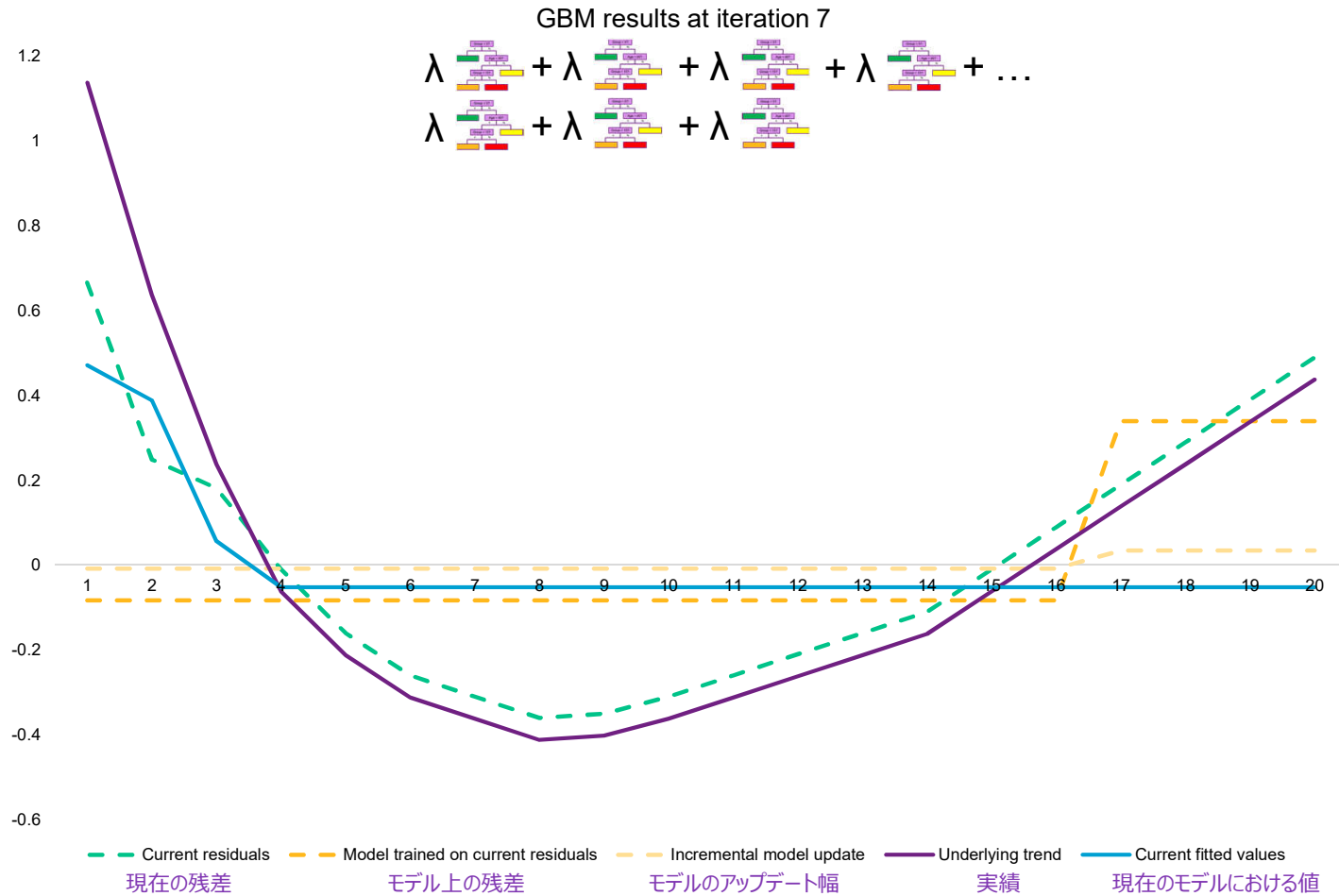
A simple GBM example (シンプルなGBMの例)



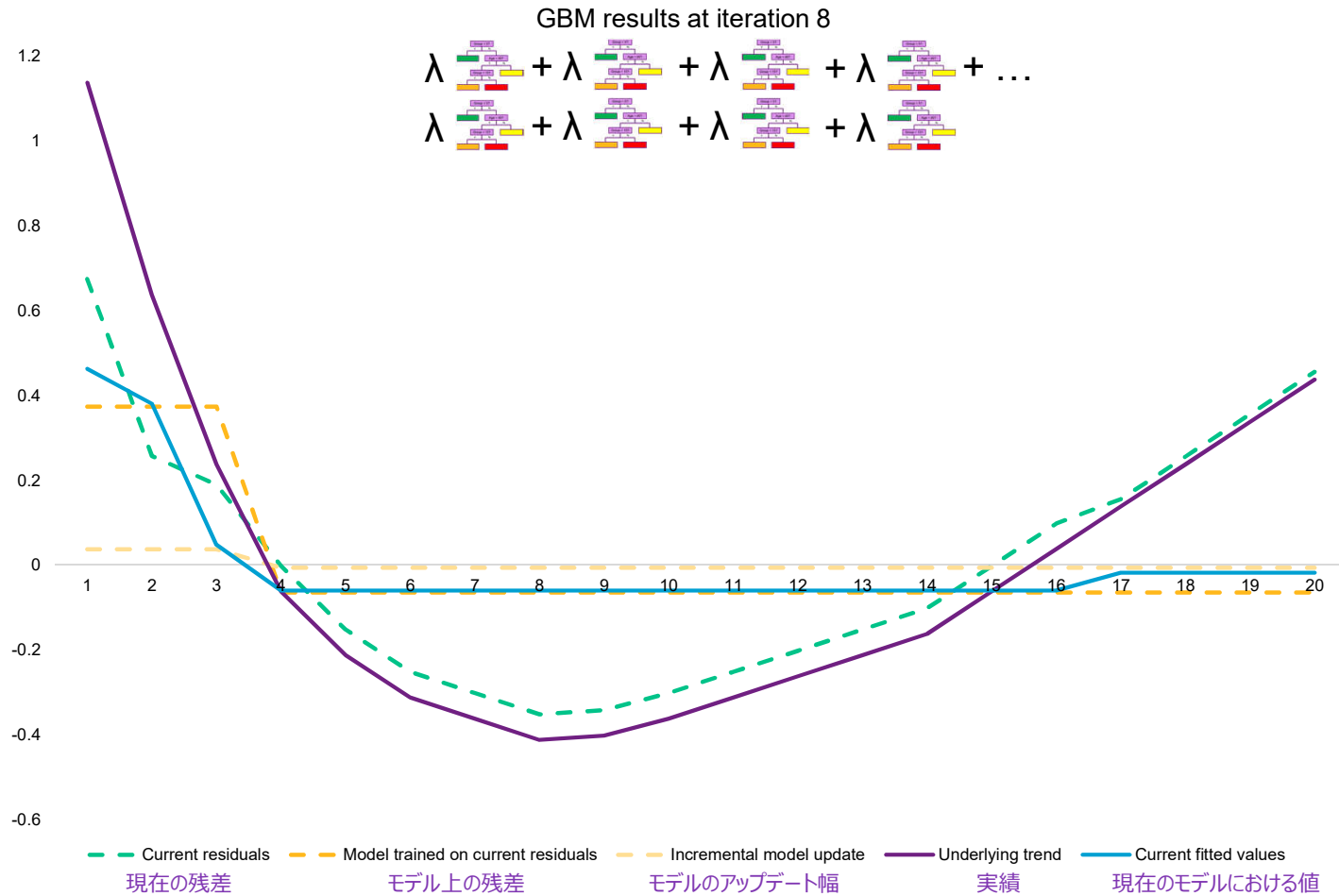
A simple GBM example (シンプルなGBMの例)



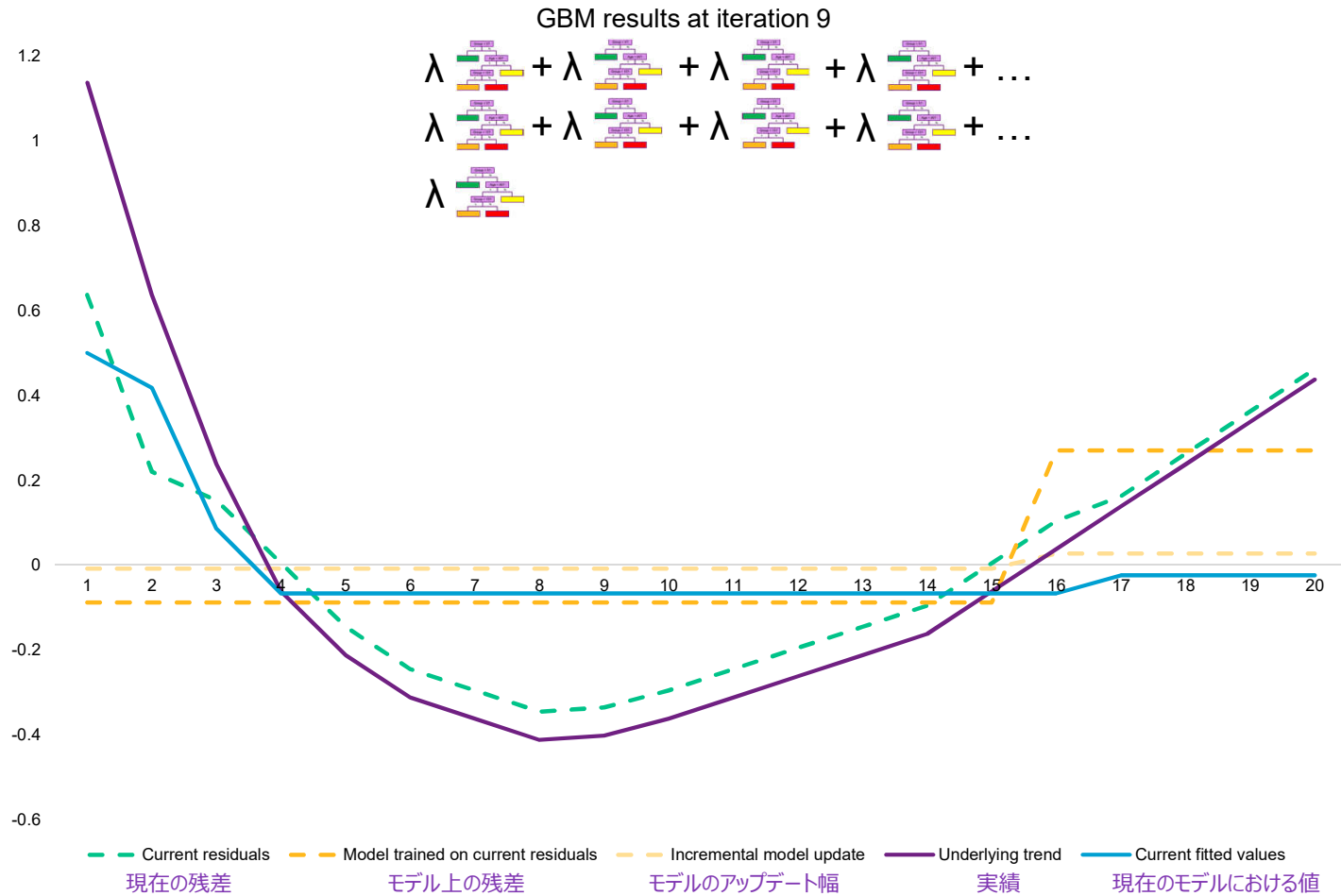
A simple GBM example (シンプルなGBMの例)



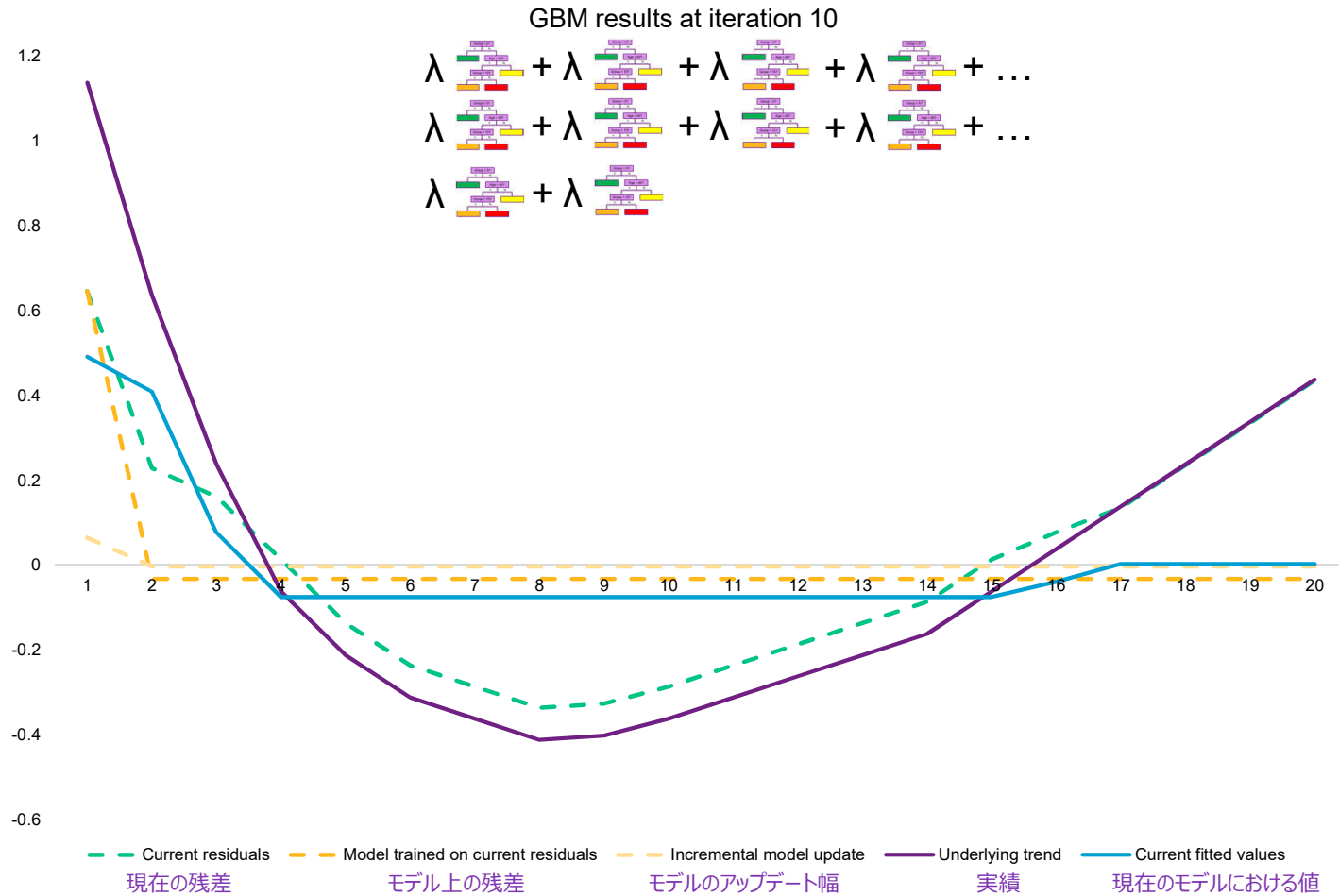
A simple GBM example (シンプルなGBMの例)



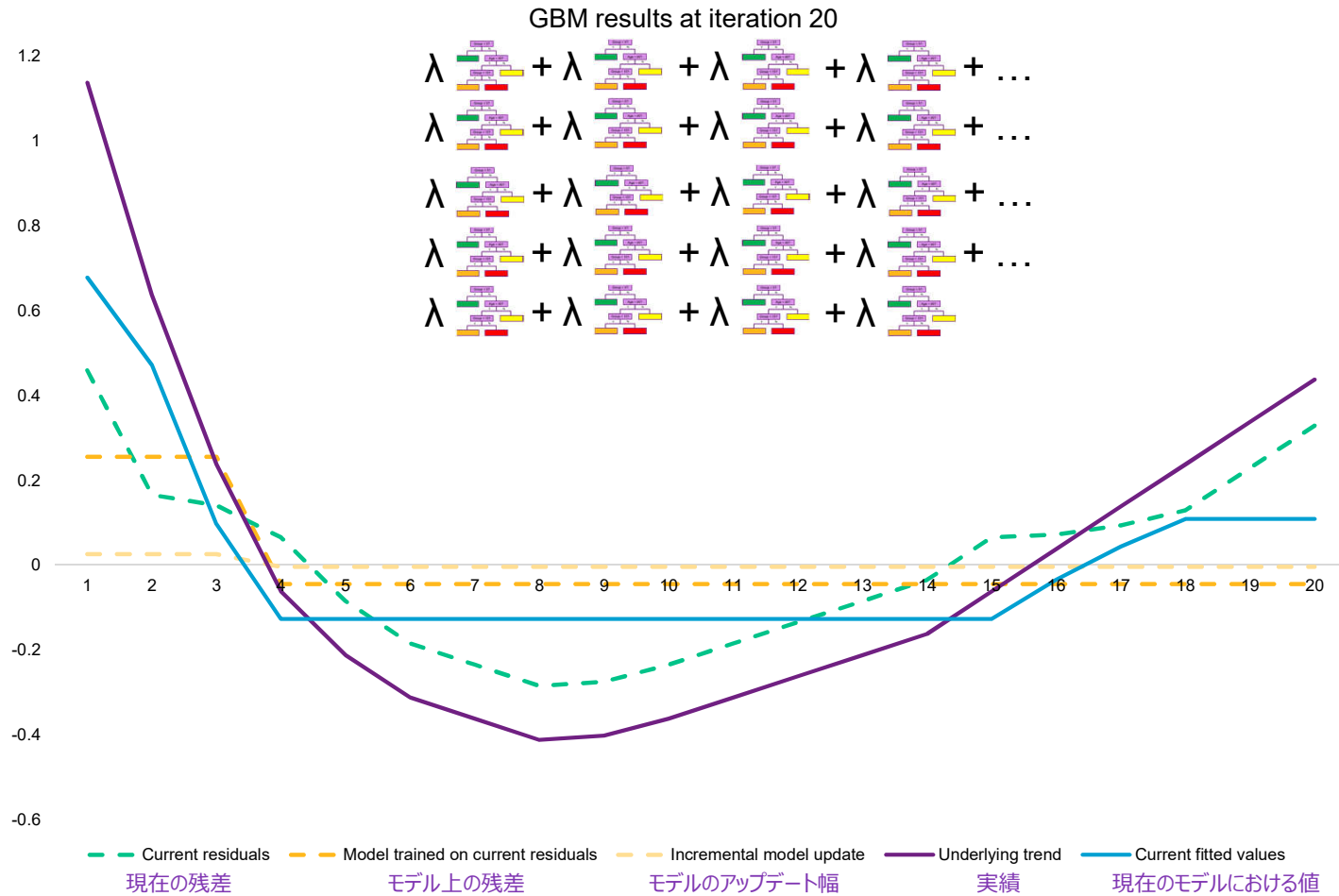
A simple GBM example (シンプルなGBMの例)



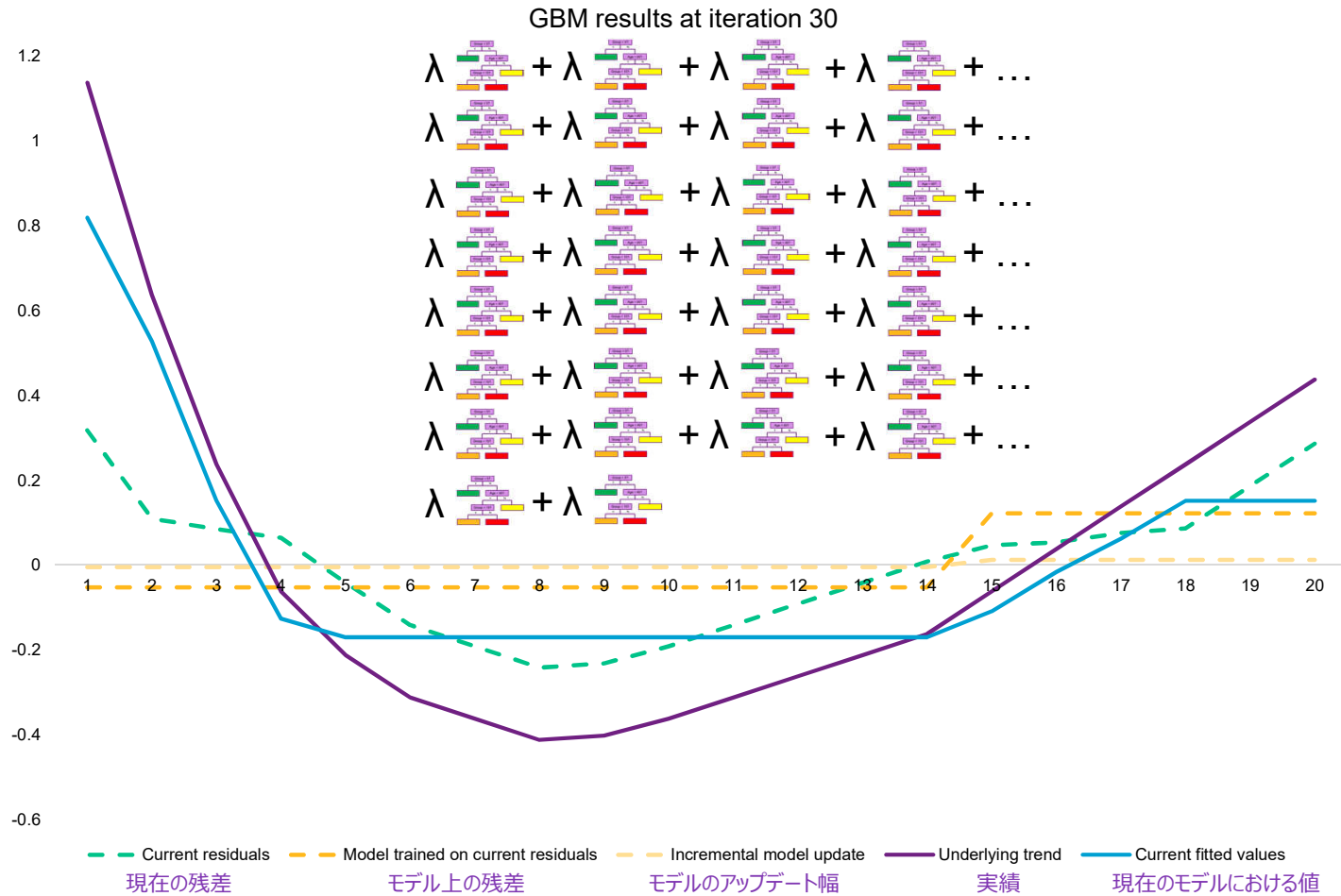
A simple GBM example (シンプルなGBMの例)



A simple GBM example (シンプルなGBMの例)

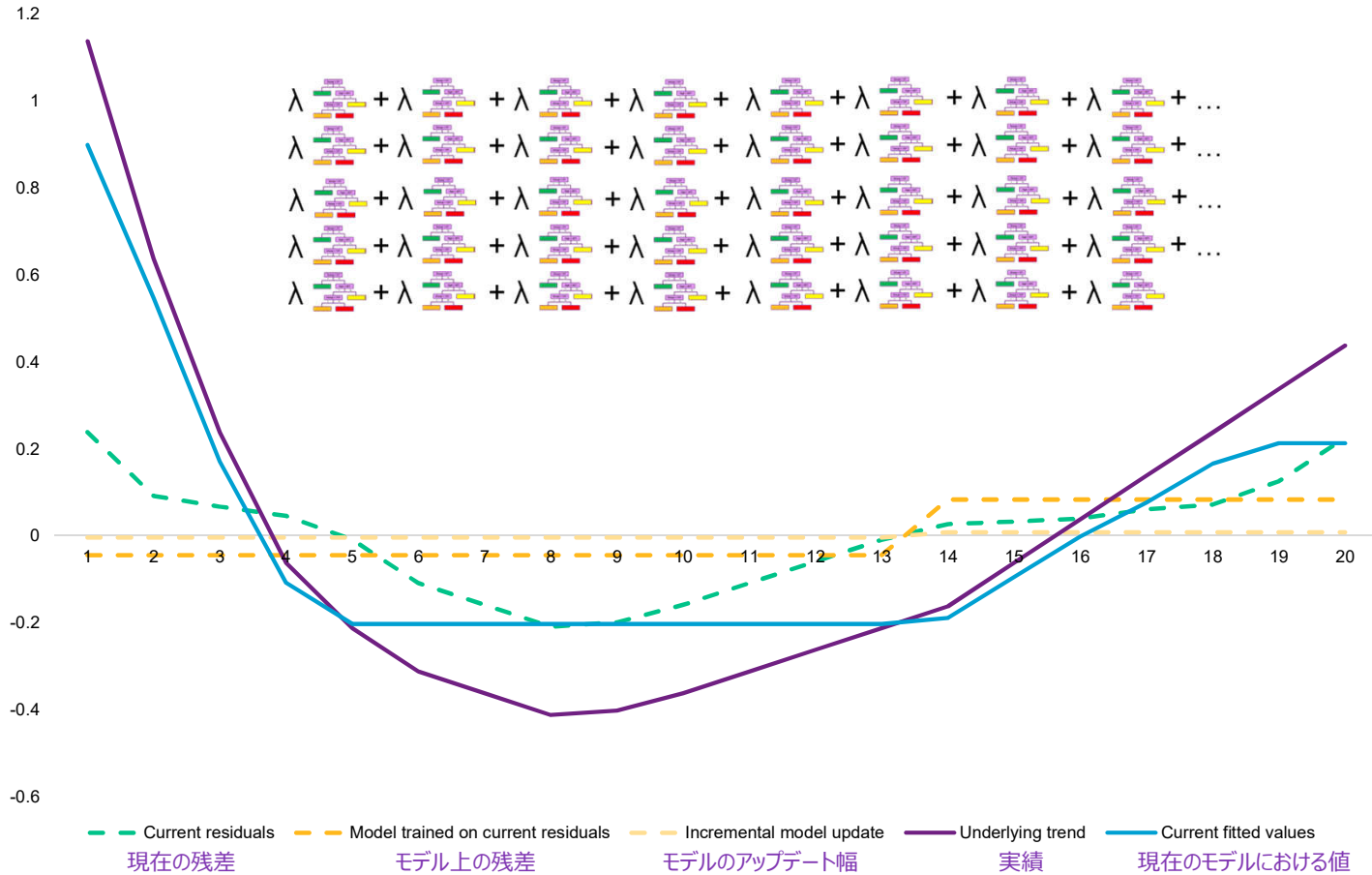


A simple GBM example (シンプルなGBMの例)



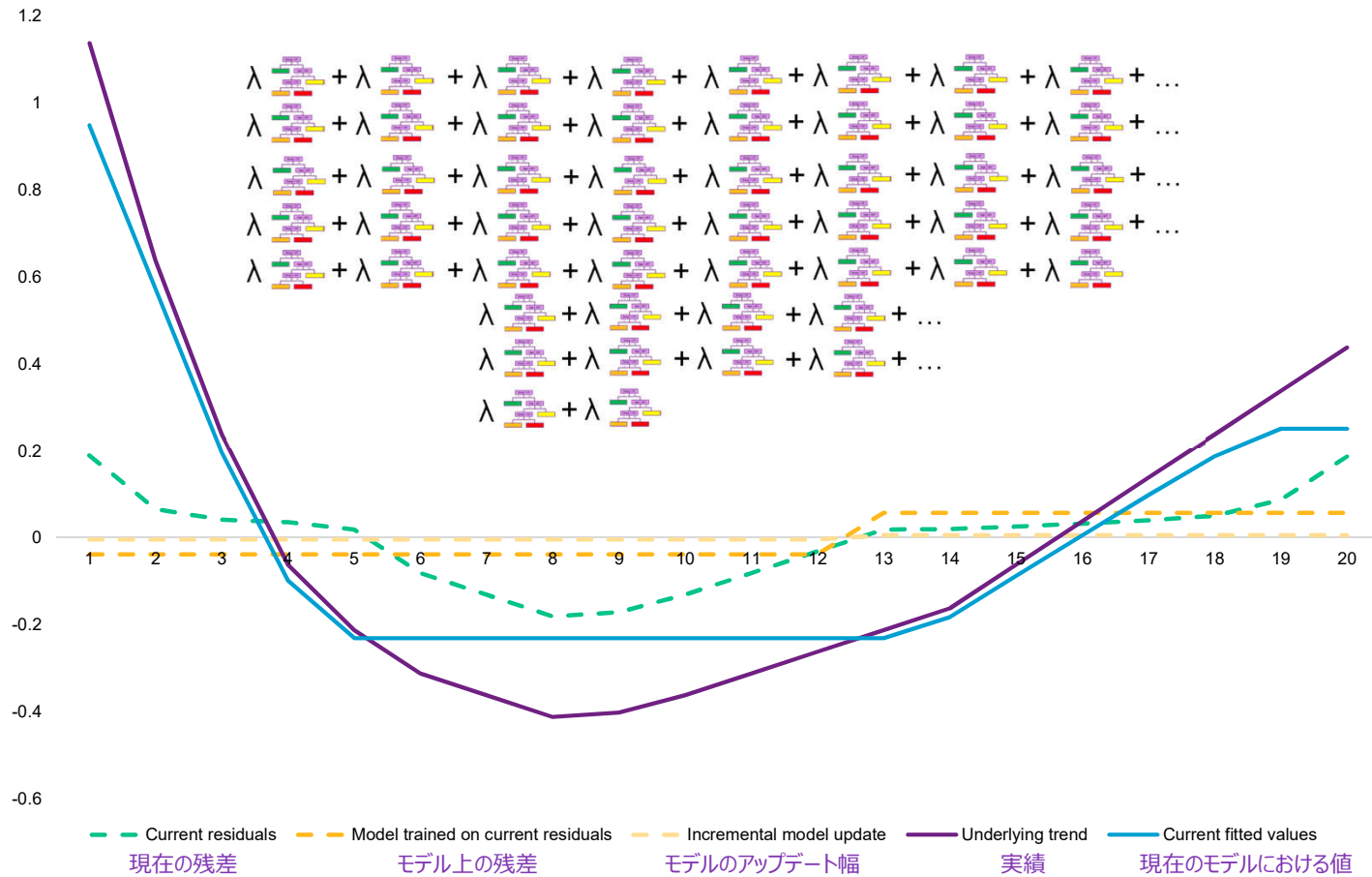
A simple GBM example (シンプルなGBMの例)

GBM results at iteration 40

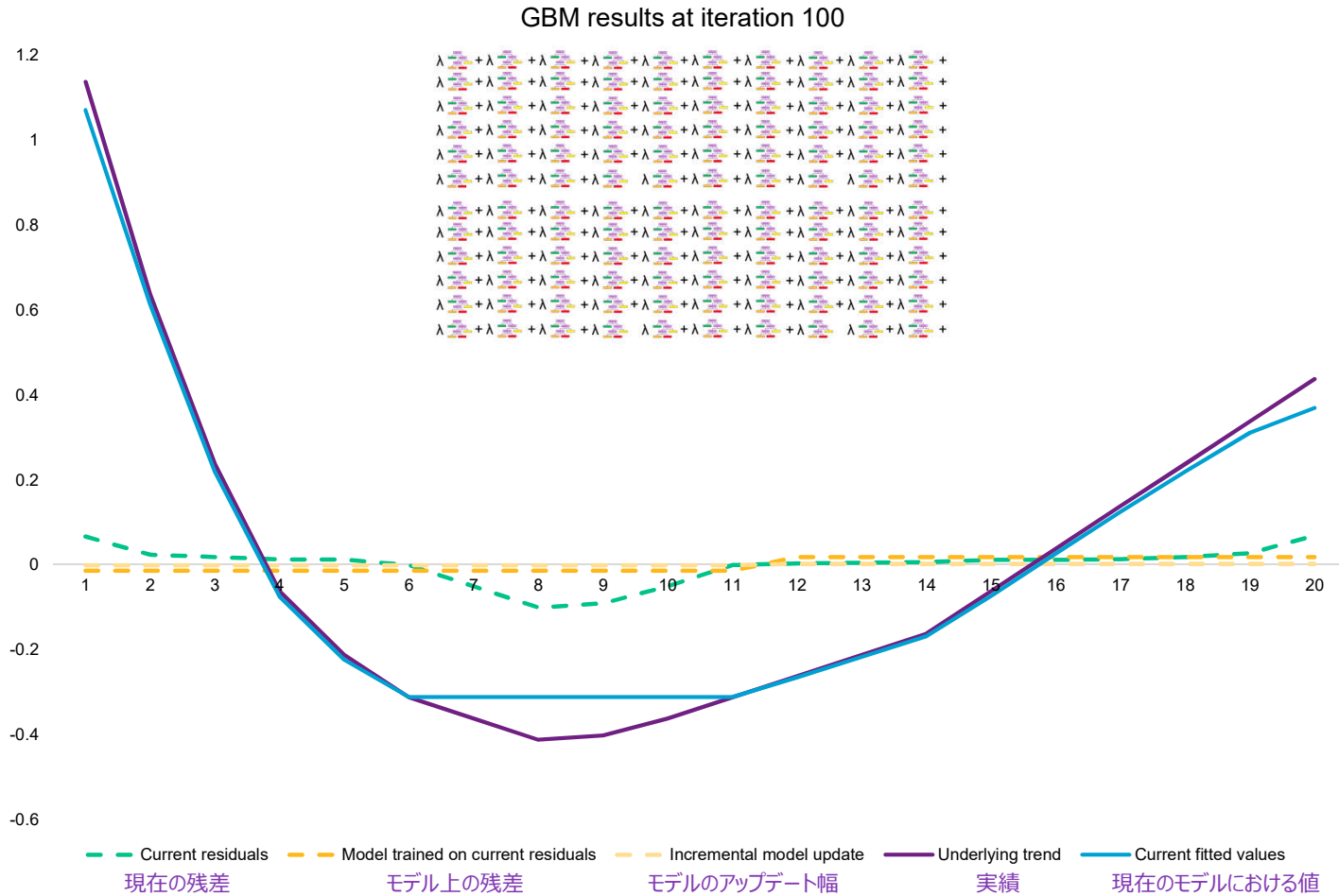


A simple GBM example (シンプルなGBMの例)

GBM results at iteration 50

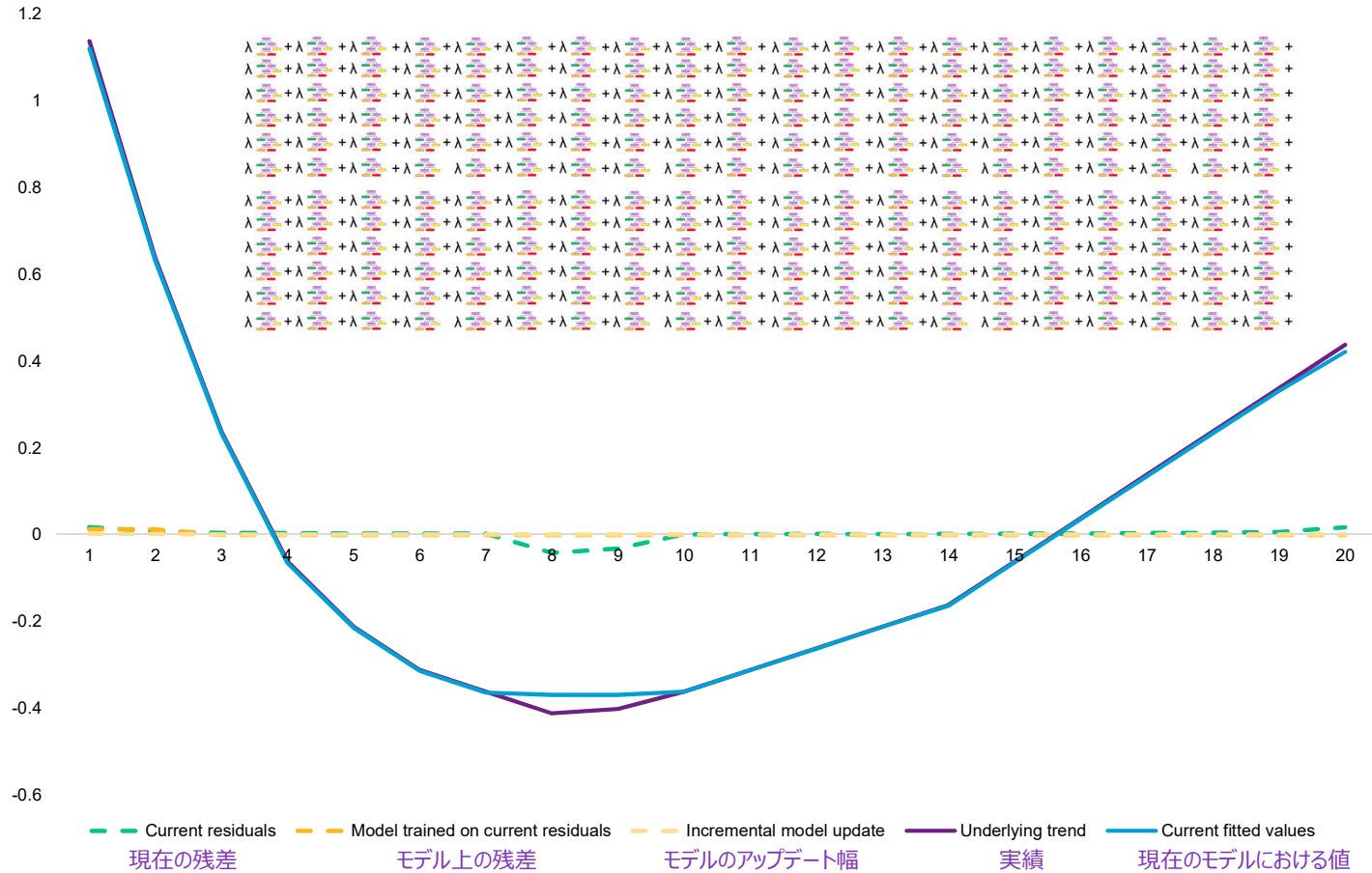


A simple GBM example (シンプルなGBMの例)



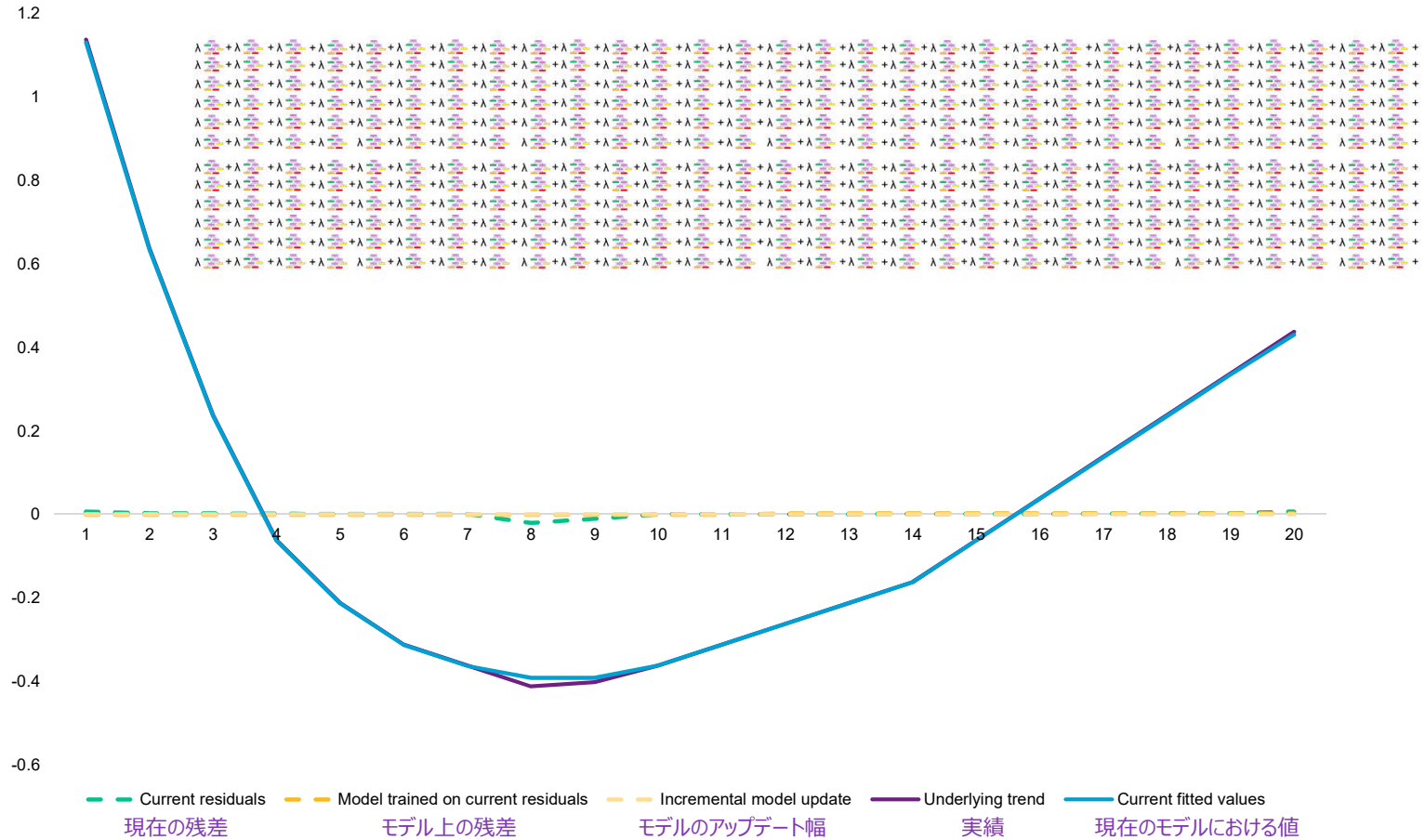
A simple GBM example (シンプルなGBMの例)

GBM results at iteration 200



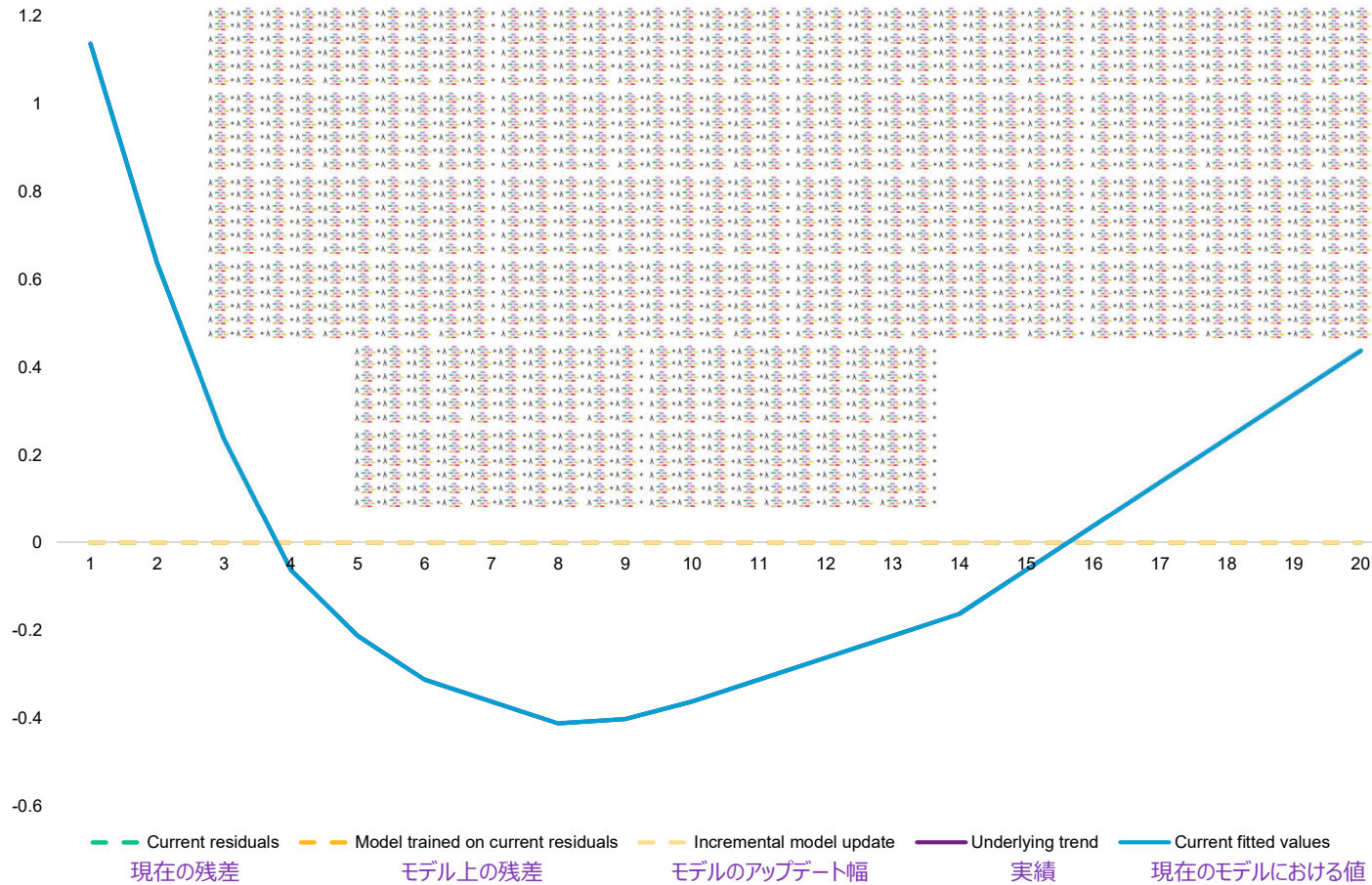
A simple GBM example (シンプルなGBMの例)

GBM results at iteration 300



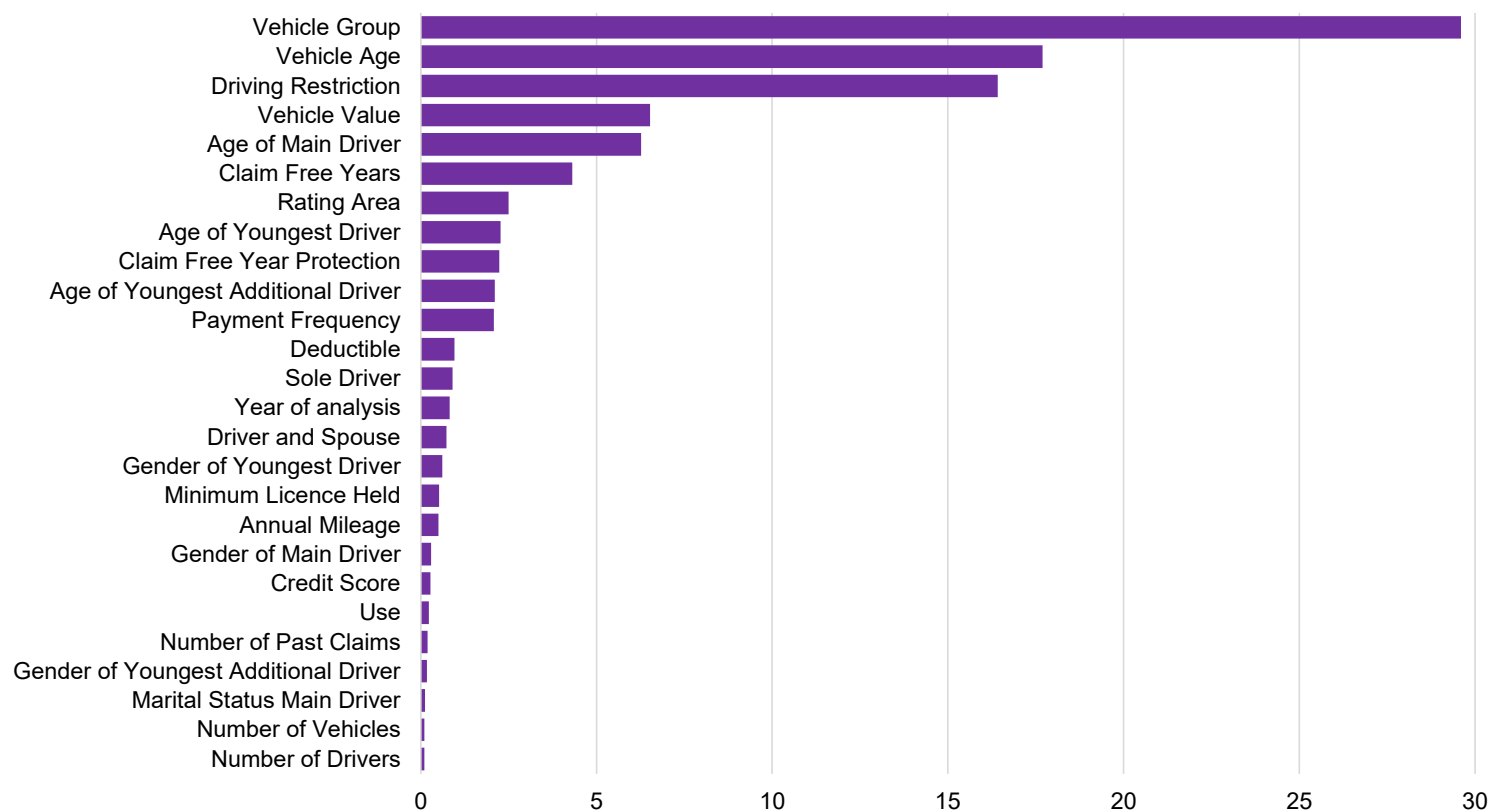
A simple GBM example (シンプルなGBMの例)

GBM results at iteration 1,000

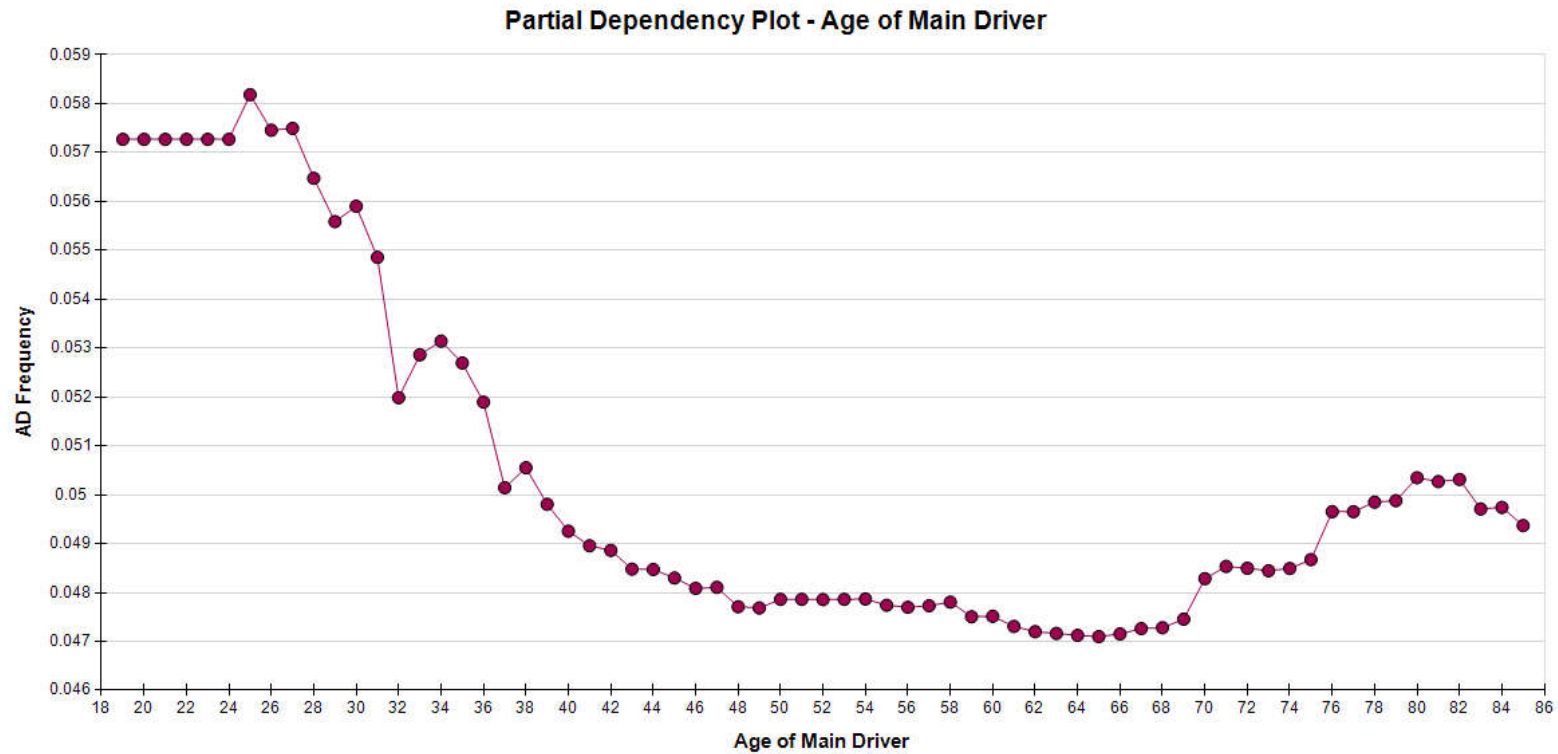


各変数の重要度 – 相対的な影響力

ファクターの重要度の違いは、GBMのすべての木に渡って、当該変数の分岐による誤差減少の合計として測定することができる。



Partial dependency plots (部分依存図)

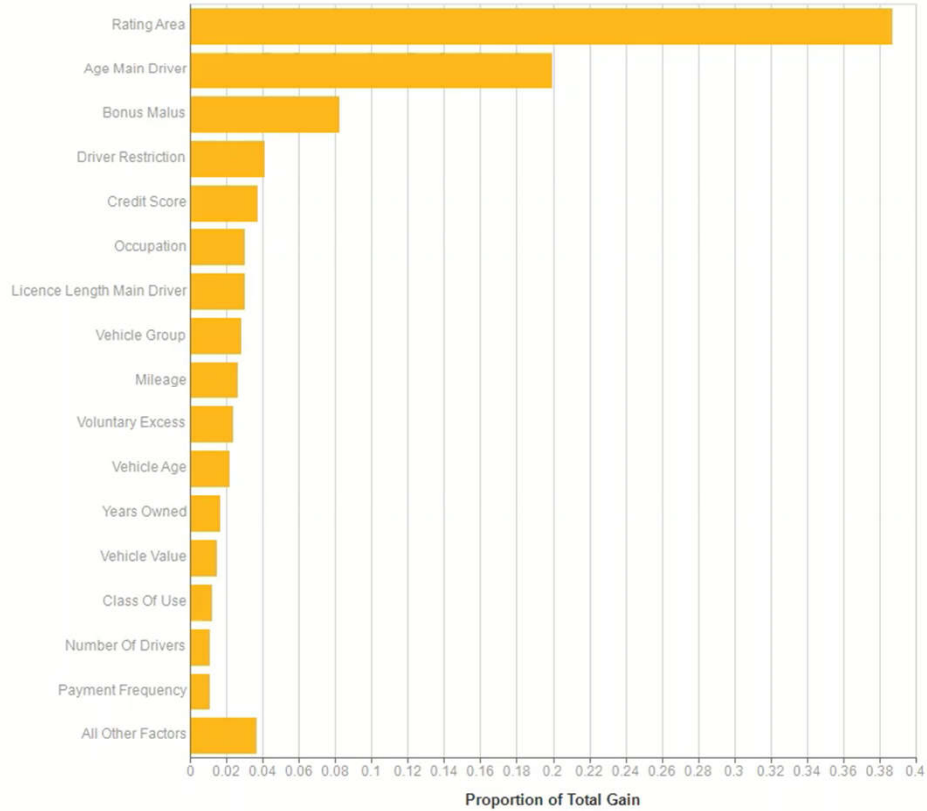


Factor Importance Selection

Bar Graph

Display All Factors Factors Per Term All Terms Show Gain From Interactions With Term

Factor Importance



GBM

Model Registry ModelRegistry (NewModel.GBMModelling.ModelRegistry)

Save Fitted Model

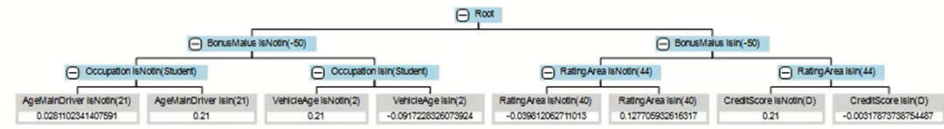
Model Analysis Item Selection

Fitted Model

GBM

Predict using trees up to: 149

Select Decision Tree 1 of 149

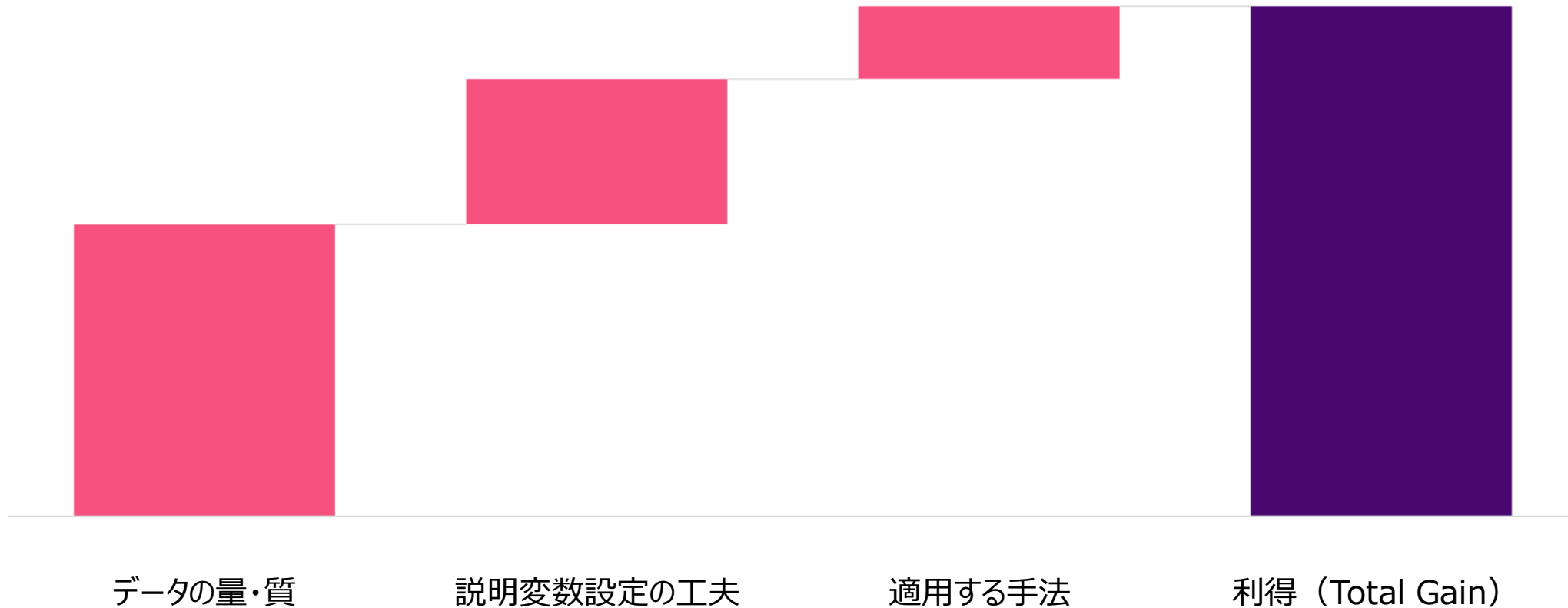


GBMs

	GLM	Penalized regression	Decision trees	GBMs
Predictive power 予測力				
Interpretation 解釈性				
Implementation 実行				
Stability 安定性				
Execution speed 実行速度				
Analytical time and effort 分析時間・ロード				

他の手法が課題を解決しうるか？

より多くのアプローチを検討することが重要



なぜ機械学習は使われるようになってきたのか？

成功のために達成すべき課題

高精度モデルの開発

- ・ 予測力、解釈可能性、安定性
- ・ モデルを用いた市場への商品投入

モデル・保険料に内在するコンプライアンス
および公正性の担保

効率性・スピードの最大化

- ・ 希少な資源の最大限の利用
- ・ マーケットニーズへタイムリーに反応する

アクチュアリー人材の専門性・技術の最大限の活用
が、ビジネスにおけるイノベーションを加速させる

プライシング・データサイエンスチームは会社の持つ最も貴重なリソース
このリソースを上手く活用できるかは、どのようにツールを利用するかにかかっている

なぜ機械学習は使われるようになってきたのか？

機械学習のメリット・デメリット

高精度モデルの開発

- ・ より高い予測力を与える機械学習
- ・ 実務導入の加速

モデル・保険料に内在するコンプライアンス
および公正性の担保
モデルの透明性に課題

効率性・スピードの最大化

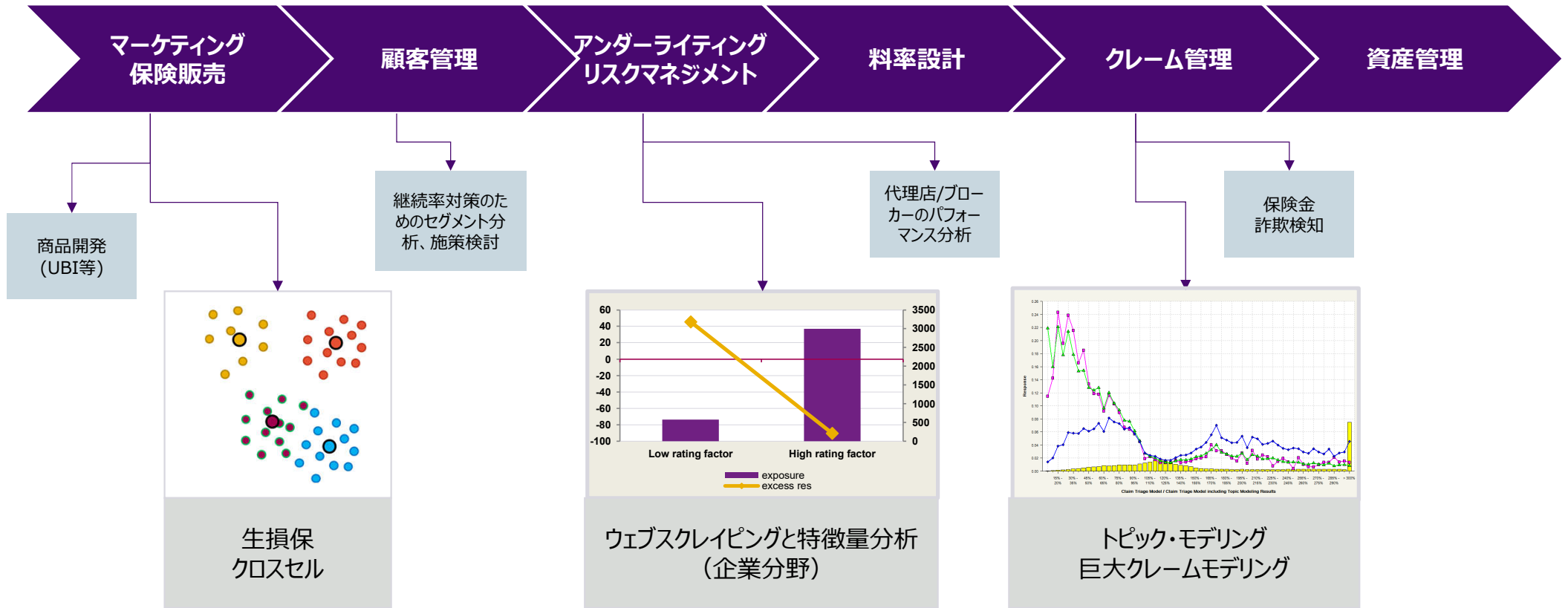
- ・ 自動化の進展
- ・ 効率性の増大

これらの技術がイノベーション（それに伴う業務プロセスの標準化）を加速させる場所に注力する必要性

プライシング・データサイエンスチームは会社の持つ最も貴重なリソース
このリソースを上手く活用できるかは、どのようにツールを利用するかにかかっている

現在 –
保険業界にて機械学習がどのよ
うに利用されているか？

保険分野における機械学習の導入



機械学習は保険事業にどの程度使用されているか？



大規模企業案件

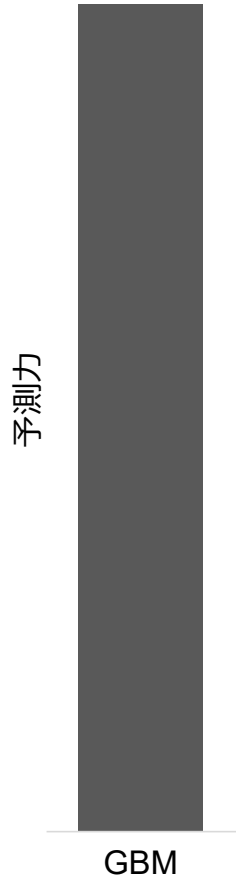
新興市場・
新しい市場

中小企業分野

米国個人商品

欧州・中東・アフリカ
個人商品

Layered GBMs

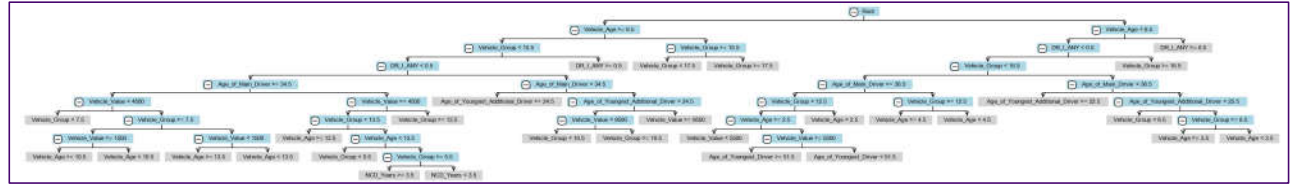


Layered GBMs

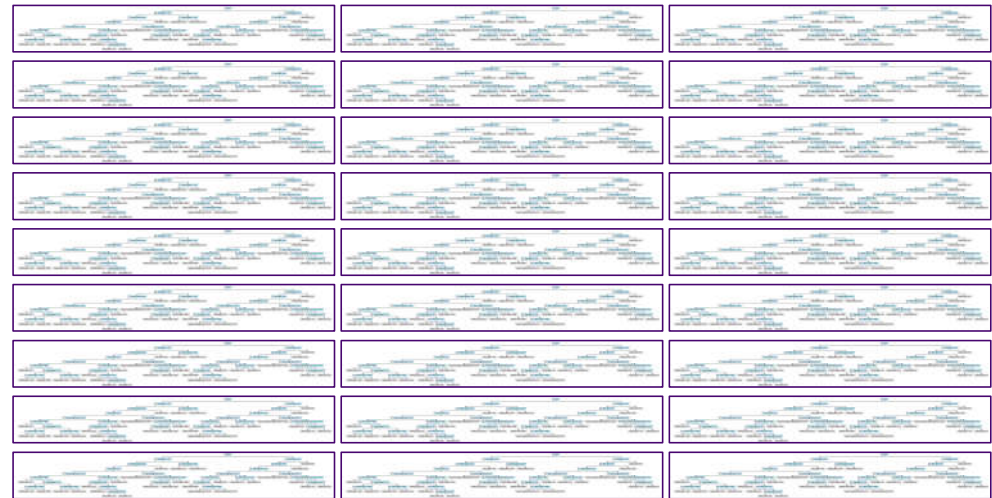
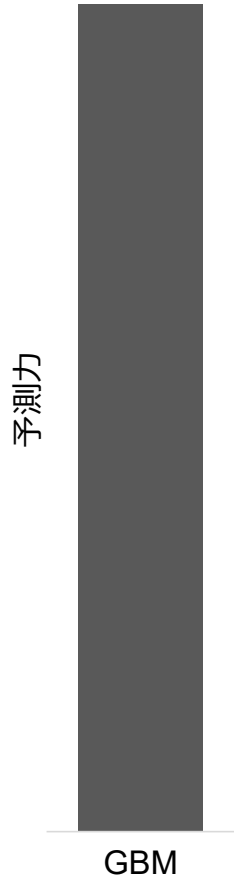
予測力



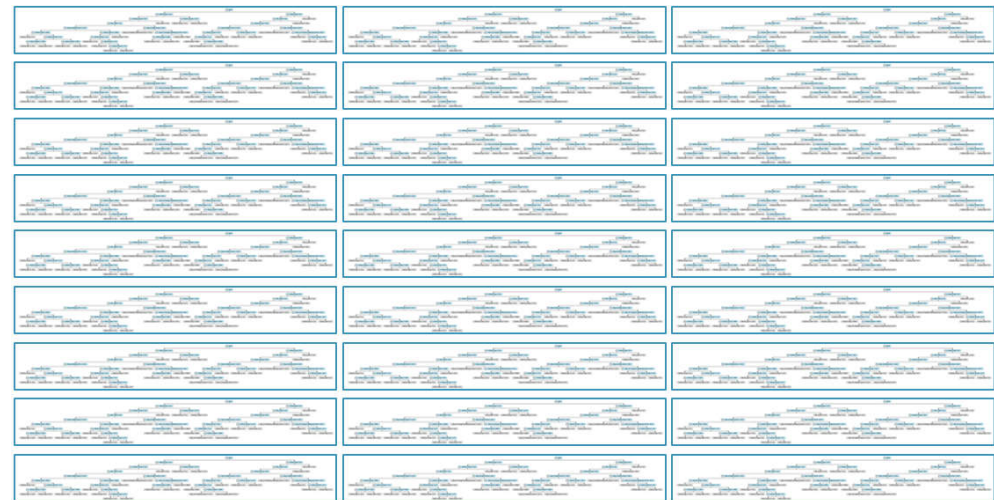
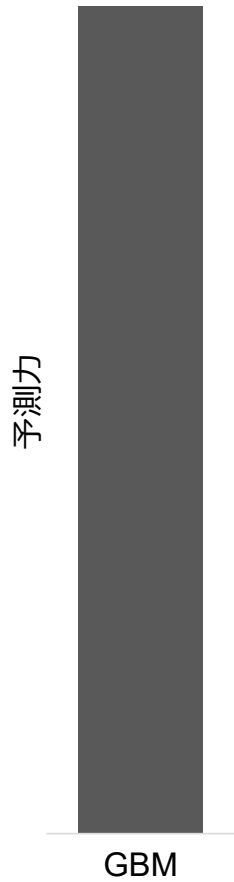
GBM



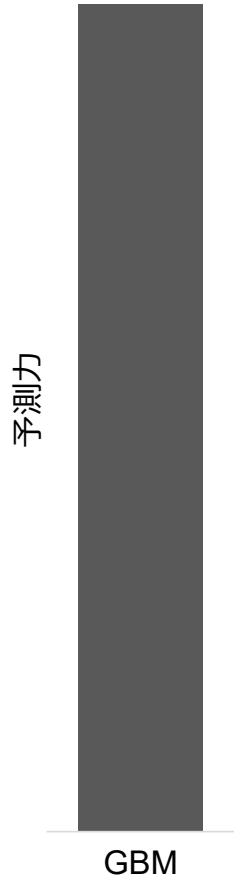
Layered GBMs



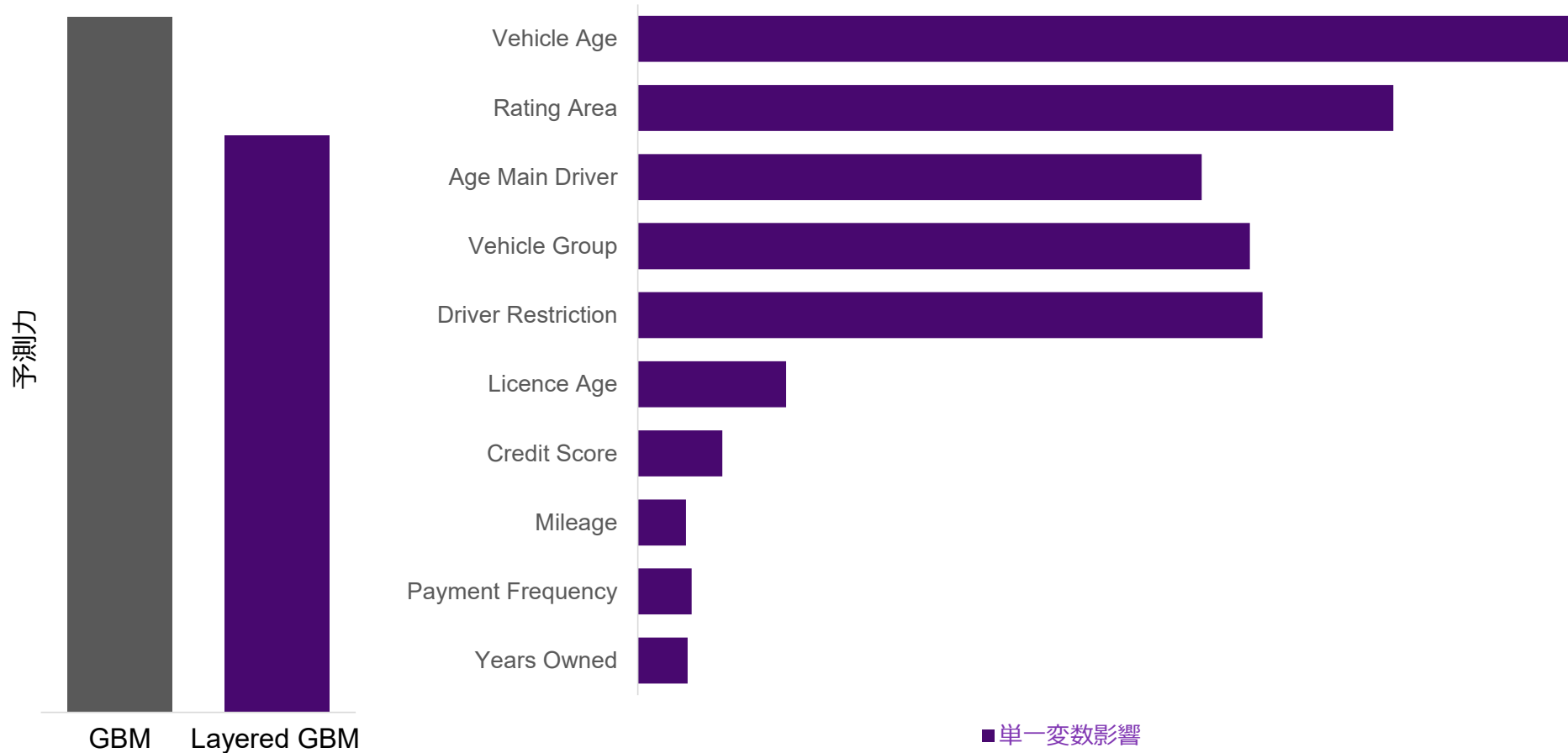
Layered GBMs



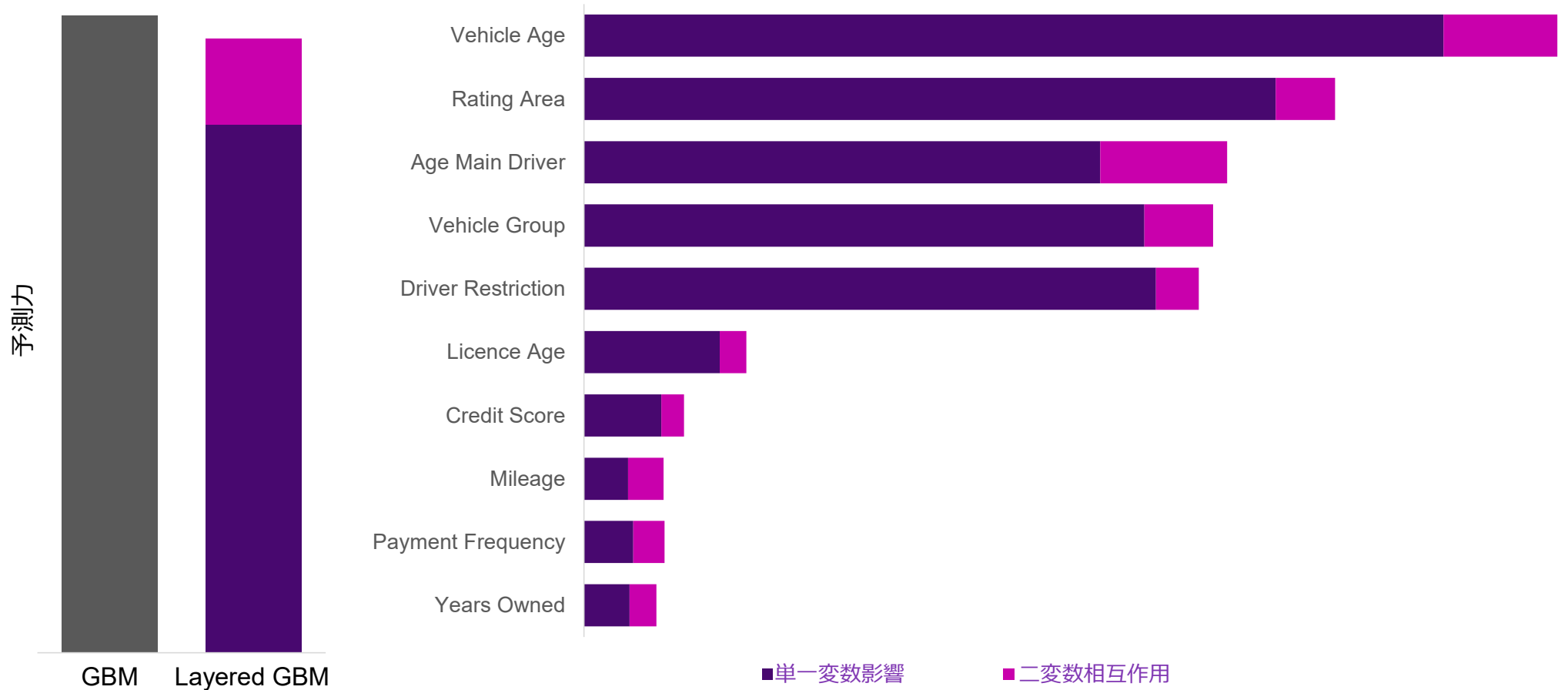
Layered GBMs



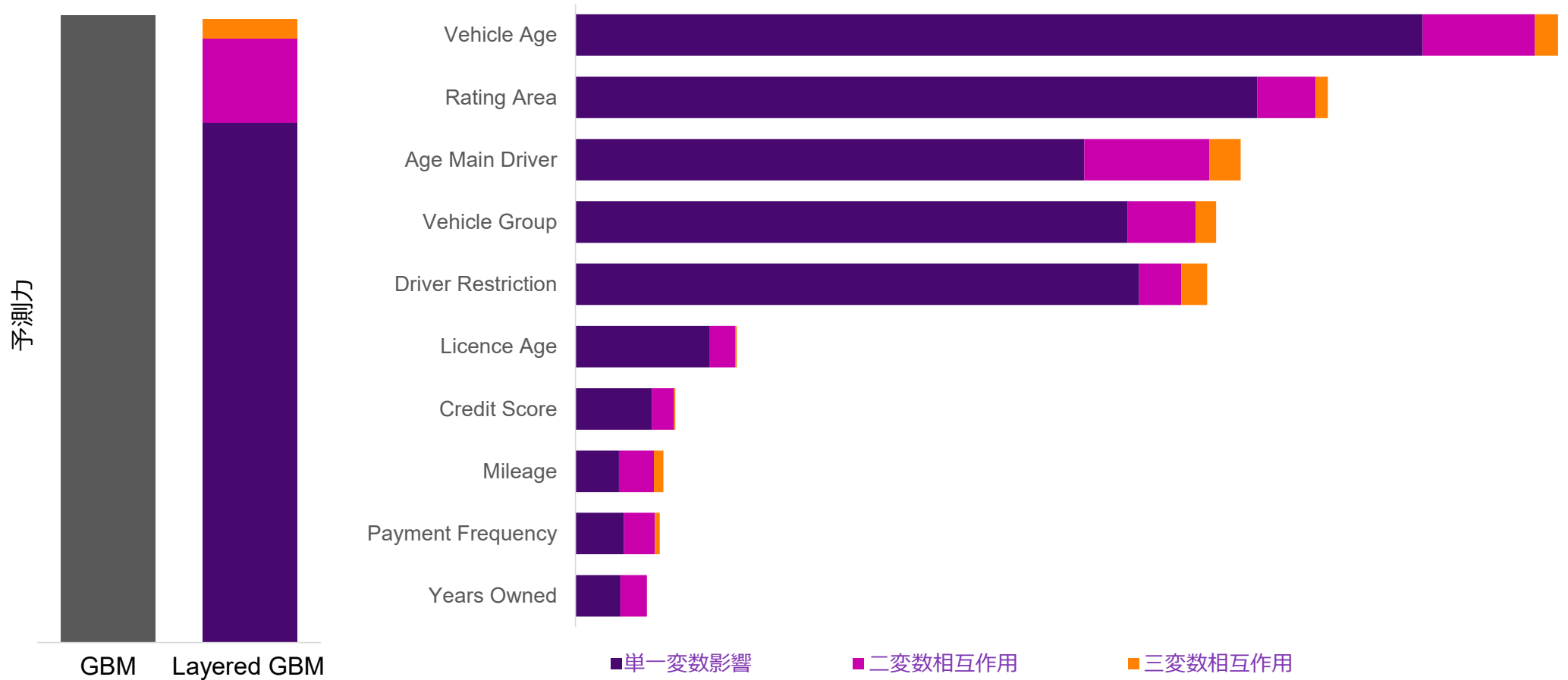
Layered GBMs



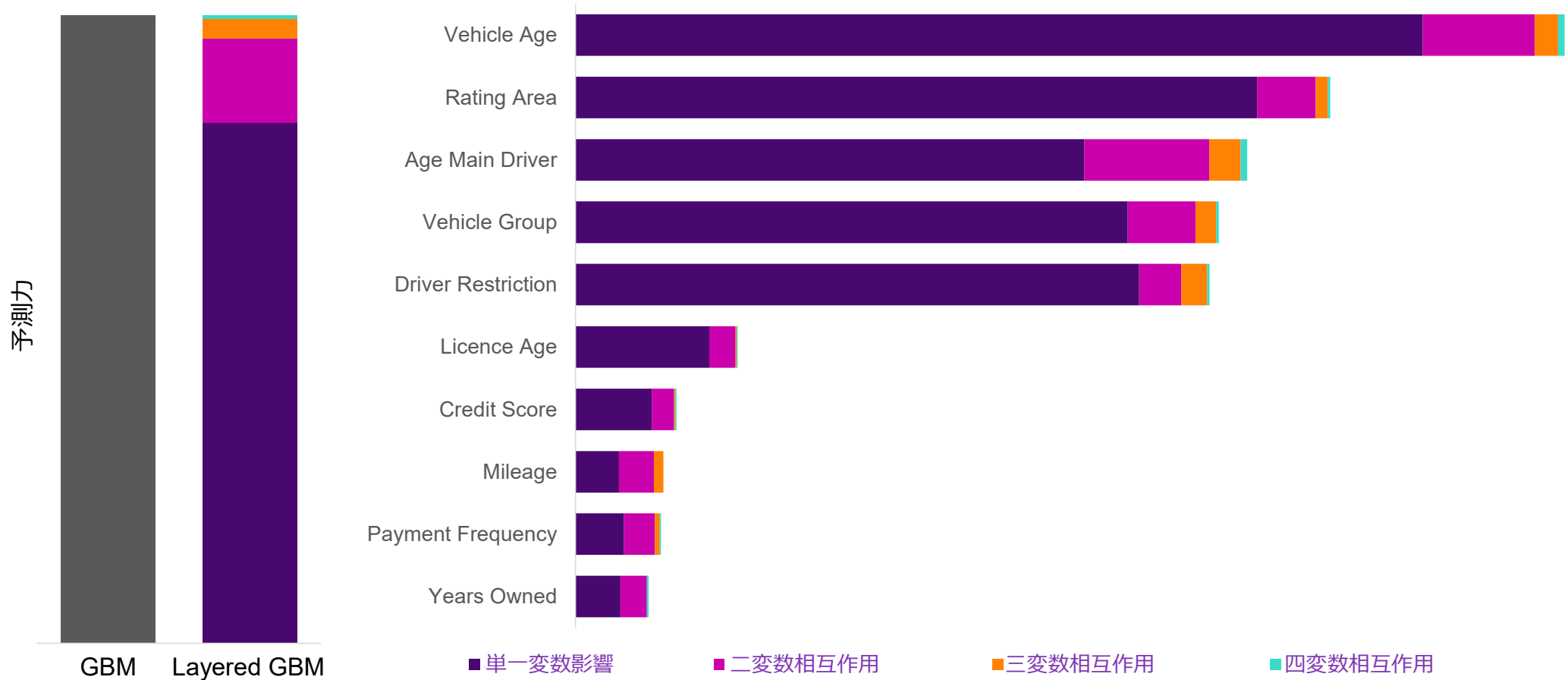
Layered GBMs



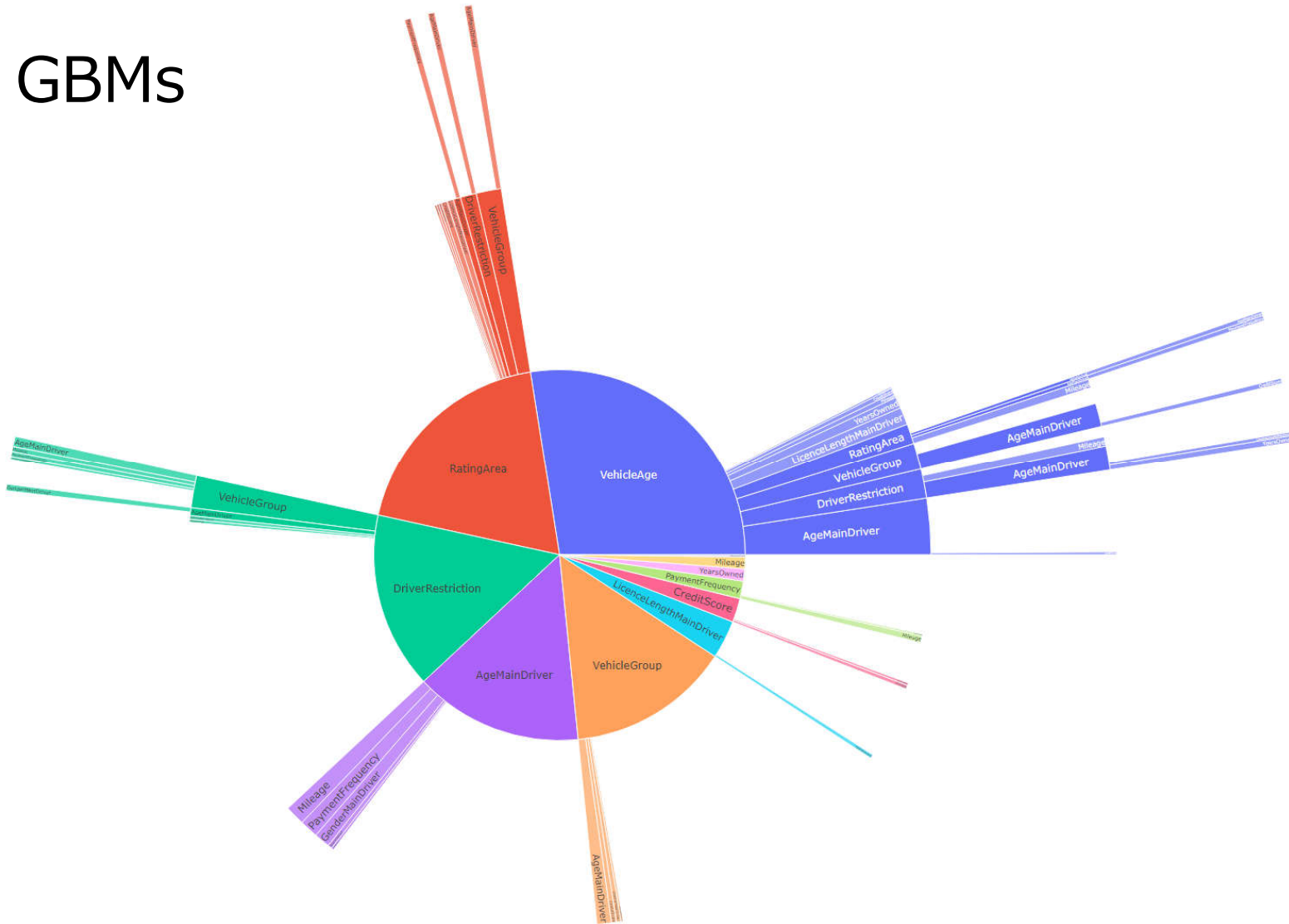
Layered GBMs



Layered GBMs



Layered GBMs

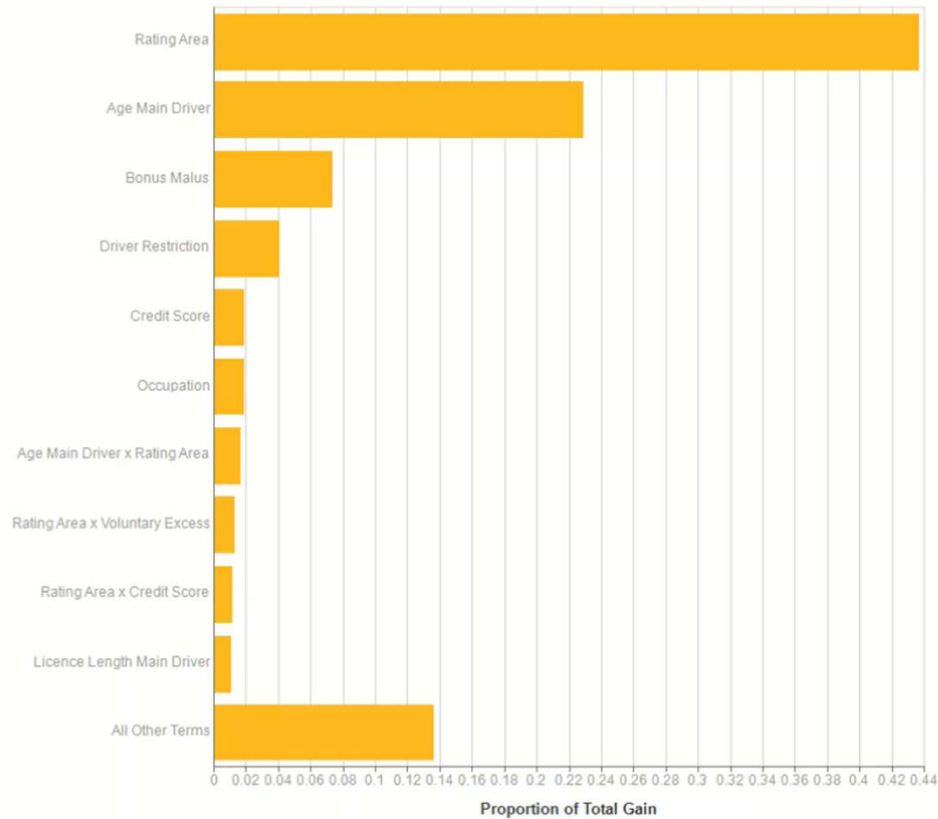


Factor Importance Selection

Bar Graph

Display All Factors Factors Per Term **All Terms** Show Gain From Interactions With Term

Factor Importance



Layered GBM
Model Registry **ModelRegistry (NewModel.LayeredGBMModelling.ModelRegistry)**
Save Fitted Model

Model Analysis Item Selection

Fitted Model

Layered GBM
Predict using trees up to: 367
Select Decision Tree 1 of 367

```
graph TD; Root["Root"] --> L["BonusMalus IsNotIn(-50)"]; Root --> R["BonusMalus IsNotIn(-50)"]; L --- LVal["0.0359679723233371"]; R --- RVal["-0.0363278234074504"];
```

Layered GBMs

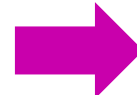


Layered GBMs

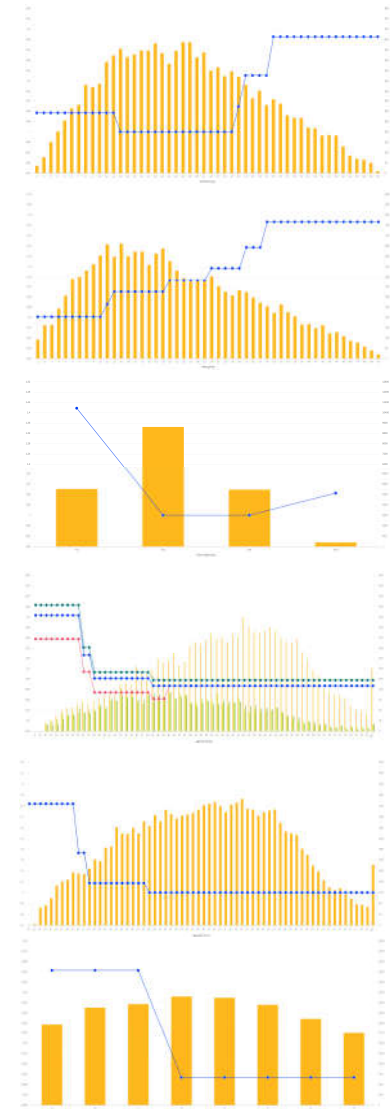
Vehicle Age	Age Main Driver x Rating Area x Vehicle Group
Rating Area	Mileage x Vehicle Age
Driver Restriction	Credit Score x Vehicle Group
Vehicle Group	Credit Score x Vehicle Age
Age Main Driver	Gender Main Driver x Vehicle Group
Licence Length Main Driver	Age Main Driver x Driver Restriction x Rating Area
Age Main Driver x Vehicle Age	Gender Main Driver
Credit Score	Age Main Driver x Driver Restriction x Gender Main Driver
Payment Frequency	Age Main Driver x Payment Frequency x Rating Area
Years Owned	Age Main Driver x Years Owned
Mileage	Licence Length Main Driver x Years Owned
Rating Area x Vehicle Group	Credit Score x Mileage
Age Main Driver x Mileage	Driver Restriction x Gender Main Driver
Age Main Driver x Payment Frequency	Age Main Driver x Driver Restriction
Licence Length Main Driver x Vehicle Age	Driver Restriction x Mileage x Vehicle Group
Driver Restriction x Vehicle Group	Driver Restriction x Payment Frequency x Vehicle Group
Age Main Driver x Gender Main Driver	Age Main Driver x Driver Restriction x Vehicle Age x Years Owned
Age Main Driver x Vehicle Group	Driver Restriction x Years Owned
Age Main Driver x Vehicle Age x Vehicle Group	Age Main Driver x Rating Area
Driver Restriction x Rating Area	Age Main Driver x Credit Score x Vehicle Age x Vehicle Group
Vehicle Age x Years Owned	Payment Frequency x Rating Area
Age Main Driver x Driver Restriction x Vehicle Age	Age Main Driver x Payment Frequency x Rating Area x Vehicle Age
Vehicle Age x Vehicle Group	Credit Score x Years Owned
Rating Area x Vehicle Age	Mileage x Rating Area
Driver Restriction x Vehicle Age	Age Main Driver x Credit Score
Age Main Driver x Driver Restriction x Vehicle Group	Age Main Driver x Rating Area x Vehicle Age x Vehicle Group
Mileage x Payment Frequency	Rating Area x Years Owned
Licence Length Main Driver x Rating Area	Payment Frequency x Vehicle Group
Driver Restriction x Mileage x Vehicle Age	Payment Frequency x Vehicle Age
Credit Score x Rating Area	Driver Restriction x Licence Length Main Driver x Vehicle Group
Mileage x Rating Area x Vehicle Age	Age Main Driver x Driver Restriction x Licence Length x Vehicle Age
	Payment Frequency x Years Owned
	Age Main Driver x Credit Score x Vehicle Age
	Driver Restriction x Licence Length Main Driver

Layered GBMs

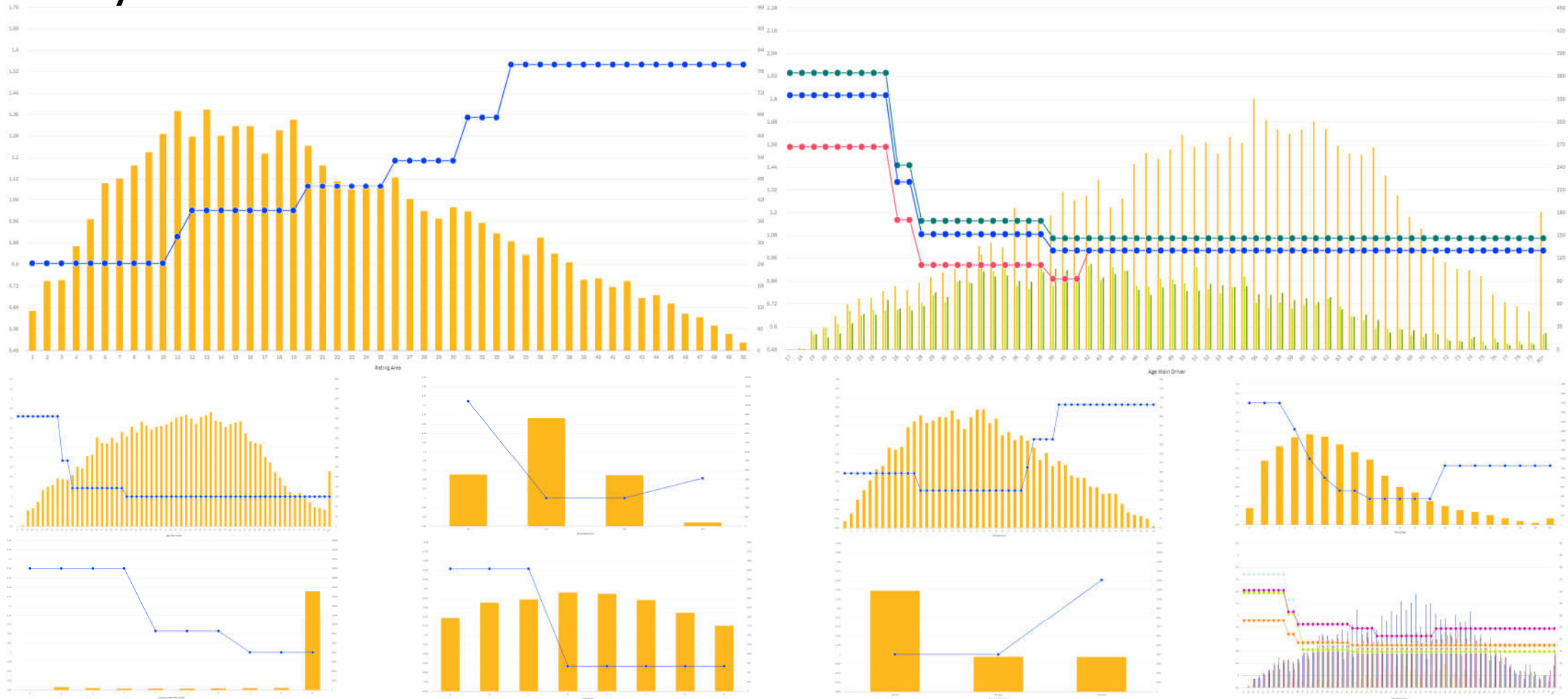
- Vehicle Age
- Rating Area
- Driver Restriction
- Vehicle Group
- Age Main Driver
- Licence Length Main Driver
- Age Main Driver x Vehicle Age
- Credit Score
- Payment Frequency
- Years Owned
- Mileage
- Rating Area x Vehicle Group
- Age Main Driver x Mileage
- Age Main Driver x Payment Frequency
- Licence Length Main Driver x Vehicle Age
- Driver Restriction x Vehicle Group
- Age Main Driver x Gender Main Driver
- Age Main Driver x Vehicle Group
- Driver Restriction x Rating Area
- Vehicle Age x Years Owned
- Vehicle Age x Vehicle Group
- Rating Area x Vehicle Age
- Driver Restriction x Vehicle Age
- Mileage x Payment Frequency
- Licence Length Main Driver x Rating Area
- Credit Score x Rating Area
- Mileage x Vehicle Age
- Credit Score x Vehicle Group
- Credit Score x Vehicle Age
- Gender Main Driver x Vehicle Group
- Gender Main Driver
- Age Main Driver x Years Owned
- Licence Length Main Driver x Years Owned
- Credit Score x Mileage
- Driver Restriction x Gender Main Driver
- Age Main Driver x Driver Restriction
- Driver Restriction x Years Owned
- Age Main Driver x Rating Area
- Payment Frequency x Rating Area
- Credit Score x Years Owned
- Mileage x Rating Area
- Age Main Driver x Credit Score
- Rating Area x Years Owned
- Payment Frequency x Vehicle Group
- Payment Frequency x Vehicle Age
- Payment Frequency x Years Owned
- Driver Restriction x Licence Length Main Driver



変数の自動単純化を
組み込んだ
Elastic Net GLM



Layered GBMs



Layered GBMs

L-GBMの利点

No loss in
predictiveness

高い予測性

Understand your
model

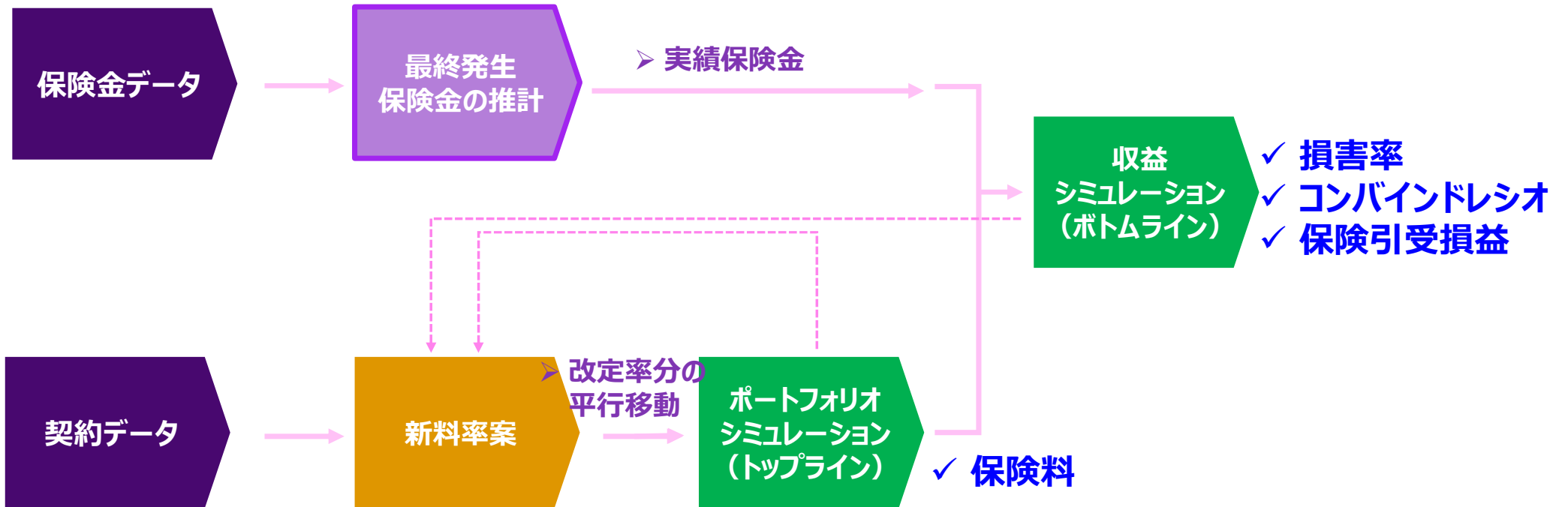
モデルの解釈

Automatic factor &
interaction detection
for GLMs

GLM用の変数選択の自動化

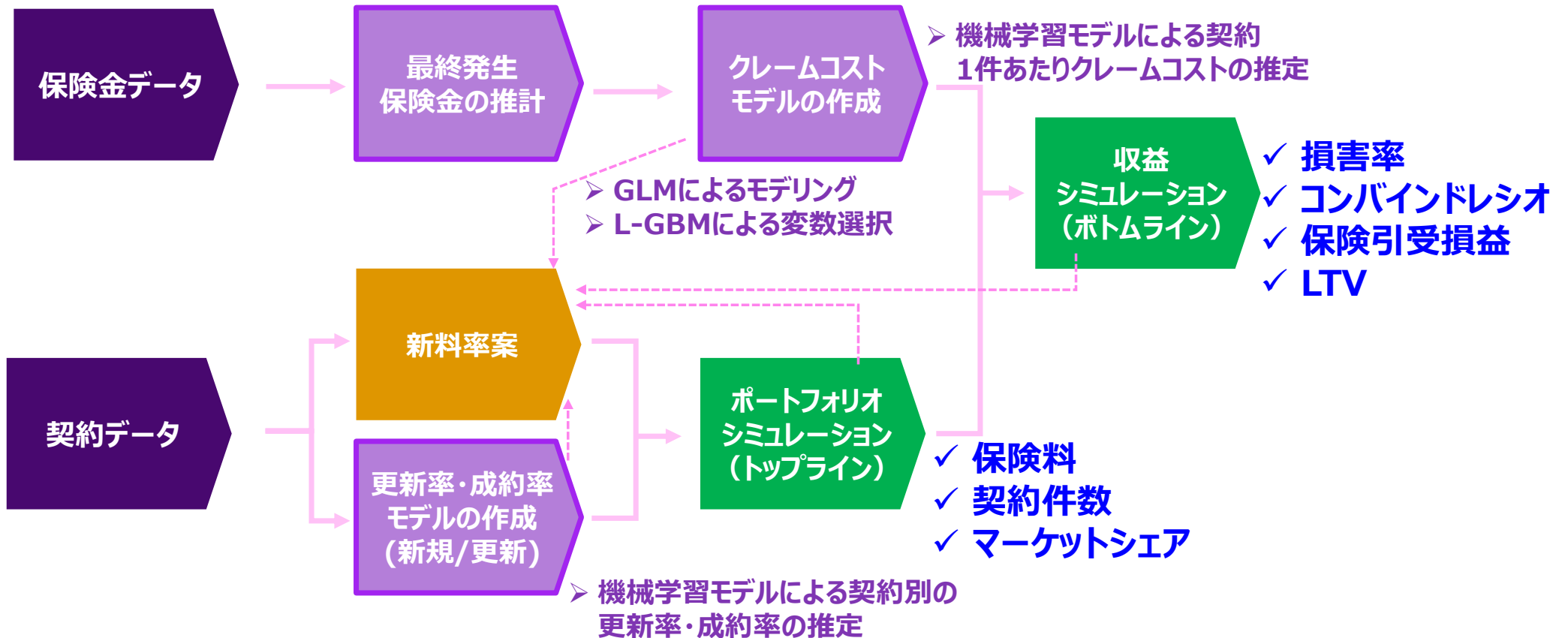
日本の保険業界における機械学習の利用

伝統的な料率算出プロセス



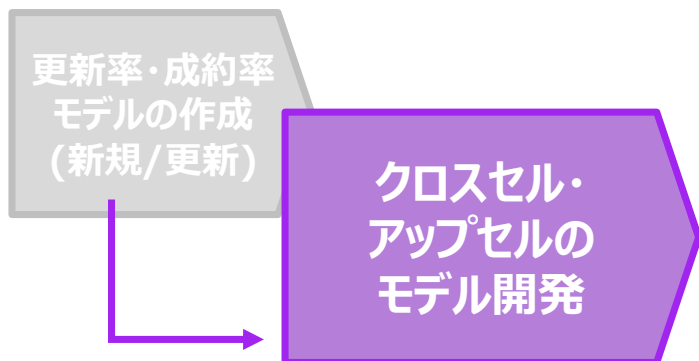
日本の保険業界における機械学習の利用

プライシングにおけるベストプラクティスの一例



日本の保険業界における機械学習の利用

プライシング以外では……



- 特定のセグメントに対する特徴の発見、施策の検討
- 営業機会の最大化



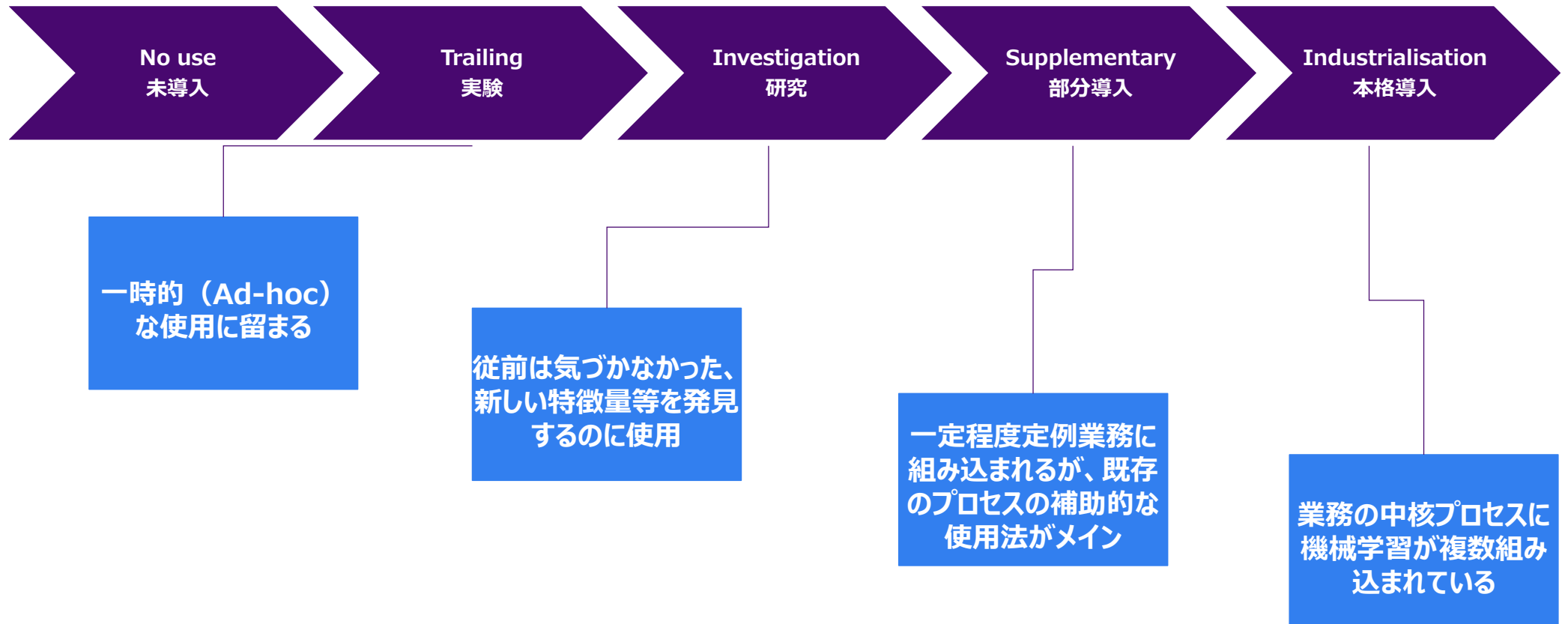
- 保険請求に関する保険金詐欺検知



- 運転データの解析によるスコアリング

etc...

保険会社の機械学習の適用、そのステップ



未来 -
今後どのように機械学習が保険
業界に貢献しうるか？

未来はどうなるのか？

Source: emeritus.org

Financial services:

詐欺検知
アルゴリズム・トレード

Healthcare:

精密医療のためのアルゴリズム
画像処理(Computer Vision)による診断・手術のサポート

Life Sciences:

製薬技術
疫病蔓延予測

Telecommunications:

ネットワーク最適化
予測メンテナンス

Oil, Gas, Energy:

画像認識のメンテナンスへの活用
エネルギー需給モデル

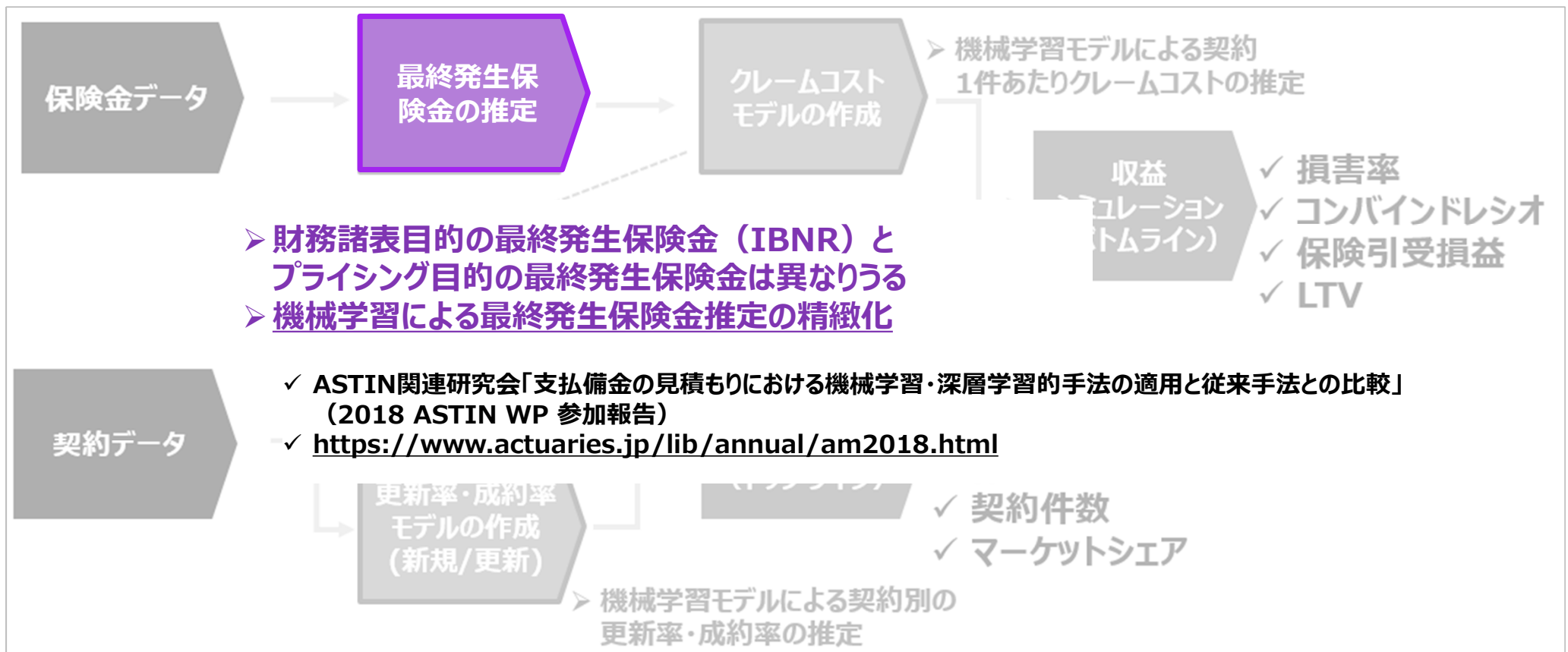
Aviation:

ルート予測モデル
顧客サービス

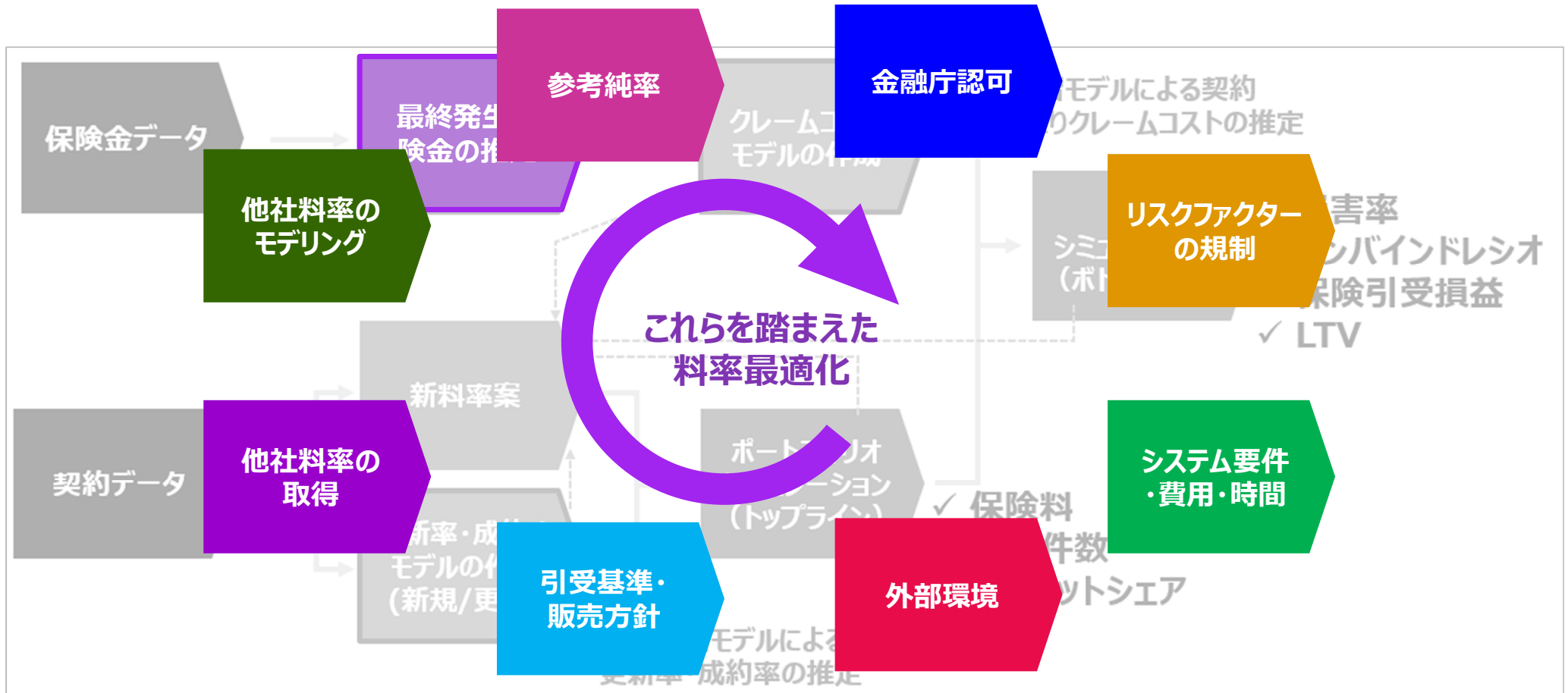
Insurance:

AIに基づくアンダーライティング
保険金請求プロセス

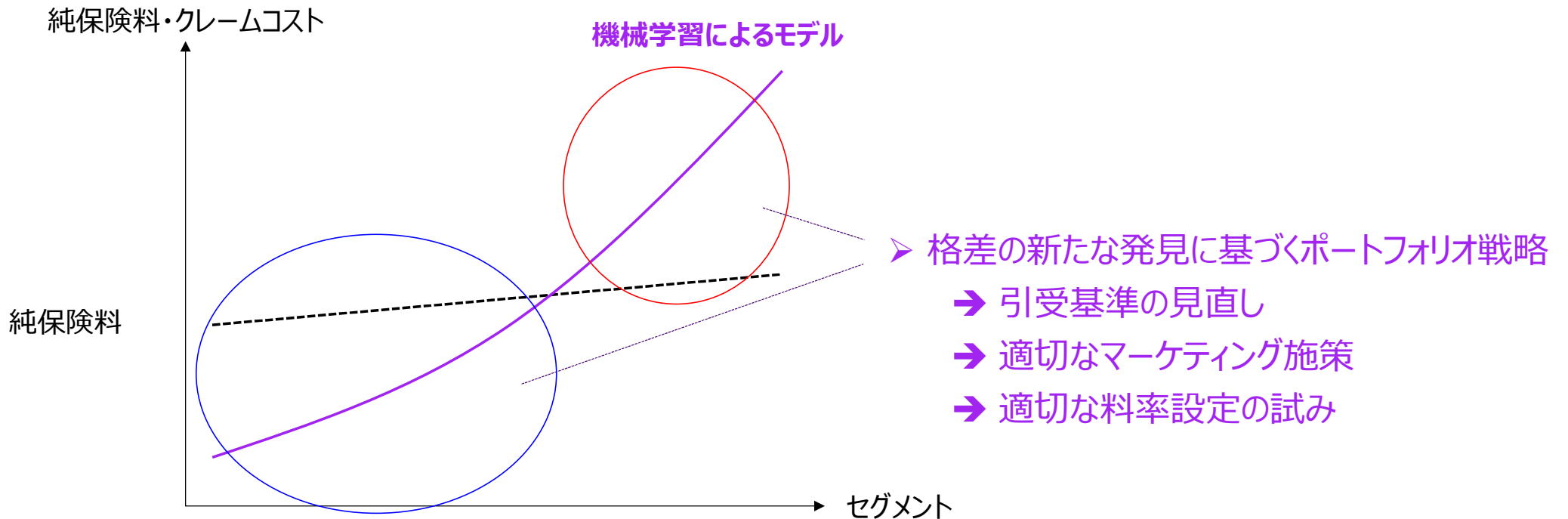
日本や他の規制の厳しい市場で、MLはどのように活用できるのか？



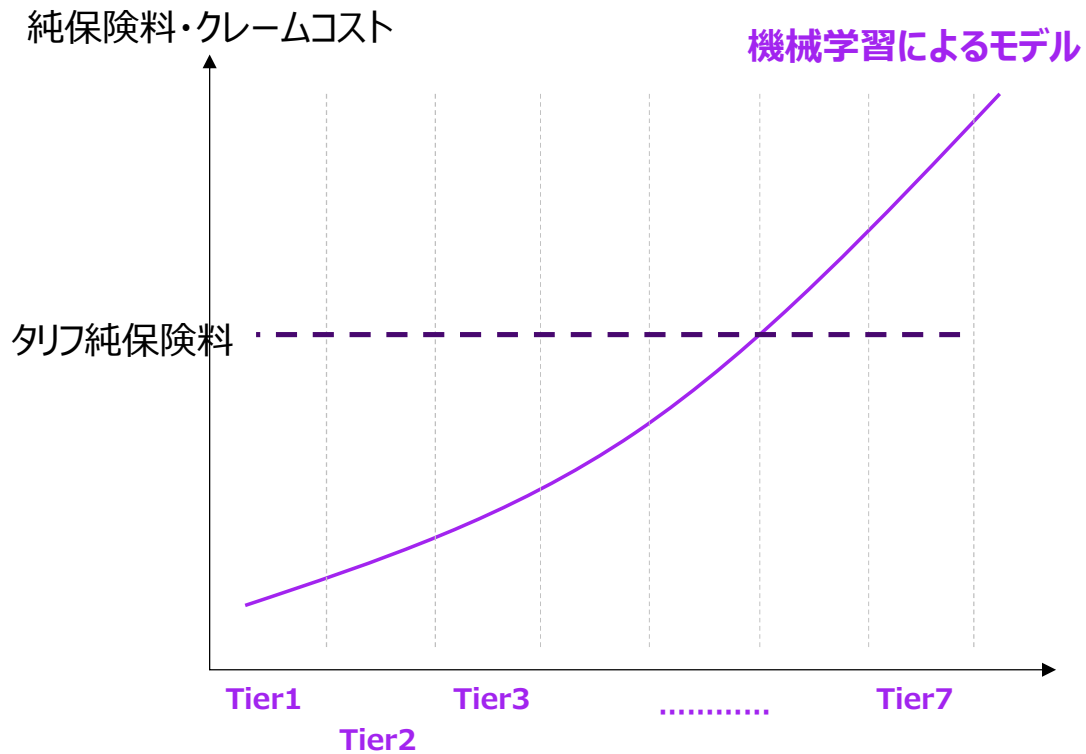
日本や他の規制の厳しい市場で、MLはどのように活用できるのか？



日本や他の規制の厳しい市場で、MLはどのように活用できるのか？



日本や他の規制の厳しい市場で、MLはどのように活用できるのか？



- 各見積（契約者）のランク付け(Tiering)
 - ➔ 区分に応じたデビエート権限の付与
 - ➔ ハイリスク区分での適切な引受

成功に向けた課題(再掲)

高精度モデルの開発

- ・ 予測力、解釈可能性、安定性
- ・ モデルを用いた市場への商品投入

モデル・保険料に内在するコンプライアンス
および公正性の担保

効率性・スピードの最大化

- ・ 希少な資源の最大限の利用
- ・ マーケットニーズへタイムリーに反応する

アクチュアリー人材の専門性・技術の最大限の活用
が、ビジネスにおけるイノベーションを加速させる

プライシング・データサイエンスチームは会社の持つ最も貴重なリソース
このリソースを上手く活用できるかは、どのようにツールを利用するかにかかっている

まとめ

- 機械学習は、保険業界の主要な解決するための大きな成長機会
- 他業界と同様に、保険業界も機械学習の最も効率的な使用方法を模索中
- 保険業界は、既存業務の中に分析業務を内包していることもあり、機械学習を活用しやすい環境
- 様々な保険マーケット・保険種目にて既に機械学習が導入されており、そこから一定の有益なアウトプットを得ている
- 日本の保険マーケットにおいてもこれらの手法を活用する明確な機会が存在しており、実際にいくつかの保険会社はそのような手法を導入している

ご清聴ありがとうございました