

倒産件数予測【モデル概要】

- 2026年Q3・Q4予測 -

伊藤海翔※1 出雲滉己※1 大里隆也※2

※1：滋賀大学データサイエンス研究科

※2：滋賀大学データサイエンスAIイノベーション研究推進センター



滋賀大学



帝国データバンク

Data Engineering and Machine Learning Center

本レポートの目的と背景

- 目的
 - 2四半期先までの業種別倒産件数を高精度に予測し、倒産動向から経済状況を可視化すること
- 背景
 - －原料価格の高騰や海外リスク、円安等により経済状況が変化していくため、**倒産件数の動向を継続的に把握する必要がある**
 - －経済状況が倒産動向に与える影響は業種ごとに異なるため、**業種別の予測が必要**である
 - －各要因が倒産件数に与える影響を把握するため、**予測精度だけでなくモデルの解釈性も重視する必要がある**

業種別倒産件数の推移

- 全体として直近の倒産件数はおおむね安定して推移
- サービス業と建設業では、**コロナ禍前の水準を上回る水準**、卸売業・小売業・製造業は**コロナ禍前と比べて同程度または低水準で推移**

業種別倒産件数の動向（2007Q1～2026Q2）^[1]



[1] シャドー部分は景気後退局面（内閣府調べ）

予測対象の業種区分について

日本標準産業分類^[1]から5つの業種区分ごとの倒産件数を集計

#	対象とする業種区分	日本標準産業分類		
		大分類	大分類名	中分類
1	卸売業	I	卸売業・小売業	50～55
2	小売業	I	卸売業・小売業	56～61
3	製造業	E	製造業	09～32
4	建設業	D	建設業	06～08
5	サービス業	L	学術研究・専門・技術サービス	71～74
		M	宿泊業・飲食サービス業	75～77
		N	生活関連サービス業・娯楽業	78～80
		Q	複合サービス業	86～87
		R	サービス業（他に分類されないもの）	88～96

- 倒産の定義についてはTDBの倒産定義のうち倒産4法を対象^[2]
- 2020年5～7月は、業種区分ごとに3ヶ月の算術平均で補正
(2020/5は、裁判所の業務縮小に伴い倒産件数が少なく見積もられているため)

[1] 総務省，日本標準産業分類，https://www.soumu.go.jp/toukei_toukatsu/index/seido/sangyo/02toukatsu01_03000044.html

[2] 帝国データバンクの倒産の定義，<https://www.tdb.co.jp/tosan/teigi.html>

予測モデルの改善実施^[1]

予測精度を維持しながら、採用変数の安定性とモデルの解釈性を向上

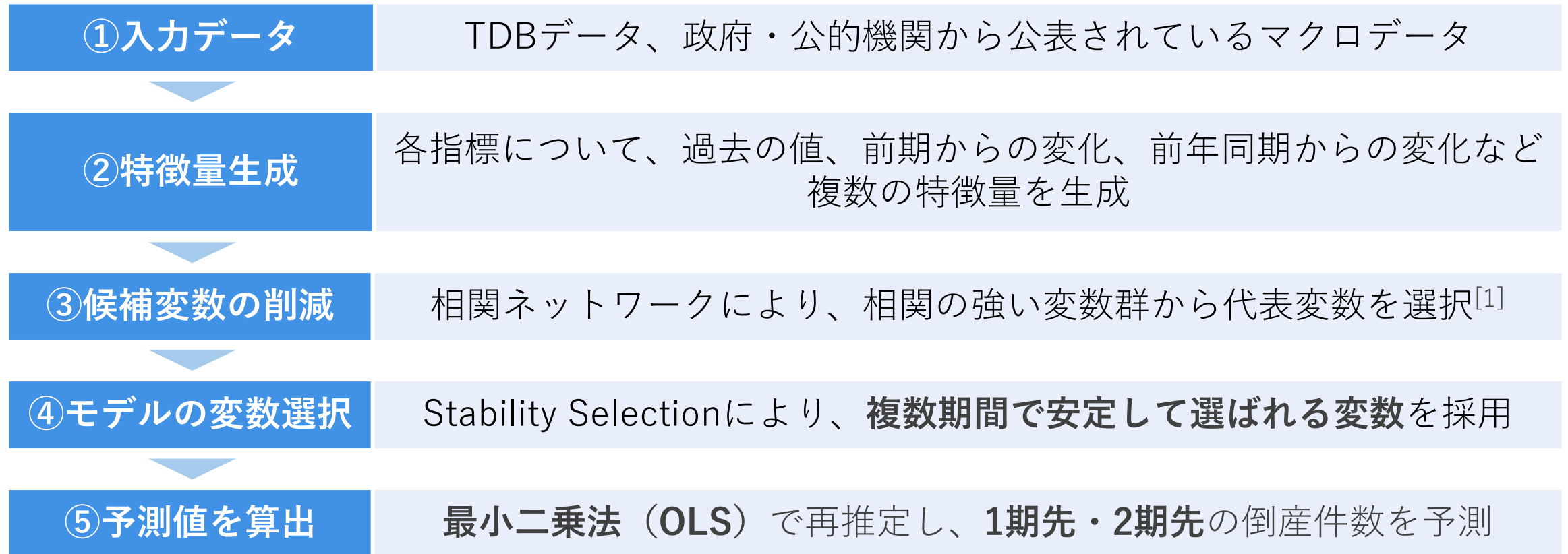
相違点	過去開発モデル（～2026Q1）	本開発モデル（2026Q2～）
目的	機械的な変数選択により直近の状況を強く反映した高精度な予測	予測精度を維持しつつ変数選択の安定性とモデルの解釈性を向上
変数選択手法	Stepwise法（変数増減法）により、四半期ごとに最適な変数を選択	Stability Selection^[2] により、複数時点で安定して選ばれる変数を採用
変数選択基準	コロナ期間を対象としたクロスバリデーション（leave-one-out交差検証法）で平均二乗誤差が最小となる変数の組合せを採用	Stability Selection での 選択確率の上位6変数 を採用 フォワードテストで予測精度と係数の安定性を確認
その他条件	変数選択を行う上で、多重共線性、不均一分散、自己相関が認められるモデルは候補から除外	変数選択を行う上で、多重共線性、不均一分散、自己相関が認められるモデルは候補から除外
モデルの特徴	コロナ禍の影響を強く反映し、四半期ごとの機械的な変数選択によるモデル更新	各四半期を通じて安定した変数を採用し、安定性、解釈性、予測精度を両立したモデルを更新

[1]過去のモデル改善についてはP.17 参考4に記載

[2]Stability Selectionの詳細はP.19 参考6に記載

予測手法の全体像

多数の説明変数候補から安定した変数を選択し、解釈可能な回帰モデルで倒産件数を予測



※モデルの詳細は次頁に記載

[1]多重共線性の排除のため、相関係数の絶対値が0.8以上の変数を削減

本開発モデルの詳細

学習データ	2012Q2～2024Q2（2020Q2は除外^[1]）を学習データ^[2]、2024Q3以降を検証データとして使用
変数選択の方法	- Stability Selectionを適用し、選択確率の高い上位6変数を候補として採用 - 候補間の相関係数の絶対値が0.8以上の場合は、選択確率が高い変数のみを残す
評価方法	2024Q4以降を最初の予測起点とし、6四半期（2四半期先予測は5四半期）のフォワードテストで予測精度とOLS係数の安定性を検証
予測モデル	1四半期先および2四半期先について、それぞれ個別にモデルを構築し、各モデルによる予測を各時点の予測値とする
変数再選択の基準	フォワードテストにおいて、予測精度もしくはOLS係数の不安定な変動がある場合、変数選択を再実施し、選択変数と予測結果を比較・検討
モデルの特徴	Stability Selectionにより、四半期ごとの予測での変数の大幅な入れ替わりを抑制し、毎四半期の予測に対して解釈性と予測の安定性、予測精度を両立 - フォワードテストで予測精度とOLS係数を確認することで、直近の経済状況と相違がないかを確認しながら運用

[1]2020Q2はコロナ期間の影響を特に強く受けているため学習から除外

[2]予測に使用可能なCCRデータが2012Q2以降に限られるため、学習期間を2012Q2以降に設定。CCRデータの詳細はP.18 参考5に記載

利用データの概要 ①

以下の説明変数^[1]を変数選択の候補とする

対象指標	説明変数となるデータの候補
倒産件数 (業種別)	倒産件数
借入環境の金利面	- 貸出約定平均金利（新規・総合・国内銀行） - 貸出約定平均金利（新規・長期・国内銀行）
対外競争力指標	実質実効為替レート
株式指標	日経平均株価_終値
業況	業況DI（中小企業・業種別）
就労状況	失業率
採算性	- 販売価格DI（中小企業・業種別） - 仕入価格DI（中小企業・業種別）
海外情勢 ^[2]	- 貿易統計 数量_PC1、数量_PC2、数量_PC3 <- 輸入品目_数量（27品目） - 貿易統計 単価_PC1、単価_PC2、単価_PC3 <- 輸入品目_単価（27品目）
民泊（サービス業）	住宅宿泊事業届出件数

[1]各データの詳細はP.13 参考1,P.14 参考2に記載

[2]貿易統計の詳細はP.15, 16 参考3-1, 3-2に記載

利用データの概要 ②

対象指標	説明変数となるデータの候補
固定費負担	- 雇用人員DI（中小企業・業種別） - 正社員数（中小企業・大企業・業種別） - 非正規社員数（中小企業・大企業・業種別） - 設備判断DI（中小企業・業種別）
借入環境の量的側面	- 資金繰りDI（中小企業・業種別） - 貸出態度DI（中小企業・業種別） - 長期借入金合計金額（中小企業・大企業・業種別） - 短期借入金合計金額（中小企業・大企業・業種別）
コロナ政策	- 日本政策金融公庫 融資実績_件数、金額 - 日本政策金融公庫 貸付残高_件数、金額 - GDP統計政府支出

- 各指標の実際の値、4期前比増減率、1期前差分を作成
- 一部指標について6か月間のボラティリティを特徴量として追加
- 1四半期先予測では、各変数の1～4四半期前の値、2四半期先予測は2～4四半期前の値を説明変数の候補とした

予測モデルの概要（1四半期先予測モデル）

変数選択の結果、1四半期先予測モデルでは以下の説明変数を用いてモデルを構築

業種	予測期間	説明変数	
サービス業	1期先	- 単価_PC3_lag4 - 新規_総合_国内銀行_4期間からの変動_lag2 - 日経平均株価_終値_4期間からの変動_lag4	- 短期借入金合計金額_サービス業_中小規模_4期間増減率_lag2 - 短期借入金合計金額_サービス業_中小規模_4期間増減率_lag4 - 短期借入金合計金額_サービス業_大規模_lag4
卸売業	1期先	- 数量_PC1_lag1 - 貸付残高_件数_4期前からの変動_lag1 - 貸付残高_金額_lag1	- 短期借入金合計金額_卸売業_大規模_lag2 - 短期借入金合計金額_卸売業_大規模_vol_6m_lag3 - 長期借入金合計金額_卸売業_大規模_lag4
小売業	1期先	- 正社員数_小売業_大規模_4期間増減率_lag4 - 短期_生産設備_中小企業_小売_差分_lag4 - 貸付残高_件数_4期間からの変動_lag4	- 貸付残高_金額_lag1 - 長期借入金合計金額_小売業_大規模_lag3
建設業	1期先	- 正社員数_建設業_中小規模_lag2 - 短期_生産設備_中小企業_建設_lag4 - 短期_貸出態度_中小企業_建設_lag1	- 長期借入金合計金額_建設業_中小規模_vol_6m_lag3 - 長期借入金合計金額_建設業_中小規模_vol_6m_lag4
製造業	1期先	- 政府支出_差分_lag4 - 数量_PC1_lag1 - 貸付残高_件数_4期間からの変動_lag2	- 長期借入金合計金額_製造業_大規模_vol_6m_lag3 - 非正規社員数_製造業_中小規模_lag2

[1] フォワードテストの結果をP.20 参考7-1に記載

予測モデルの概要（2四半期先予測モデル）

変数選択の結果、2四半期先予測モデルでは以下の説明変数を用いてモデルを構築^[1]

業種	予測期間	説明変数	
サービス業	2期先	- 単価_PC3_lag4 - 新規_総合_国内銀行_4期前からの変動_lag2 - 日経平均株価_終値_4期前からの変動_lag4	- 短観_貸出態度_中小企業_宿泊_飲食サービス_差分_lag2 - 短期借入金合計金額_サービス業_中小規模_4期前増減率_lag4 - 短期借入金合計金額_サービス業_大規模_lag4
卸売業	2期先	- 単価_PC3_4期前からの変動_lag4 - 失業率_差分_lag2 - 数量_PC1_lag4	- 新規_総合_国内銀行_vol_6m_lag4 - 短観_貸出態度_中小企業_卸売_差分_lag2 - 貸付残高_件数_4期前からの変動_lag3
小売業	2期先	- 数量_PC3_4期前からの変動_lag3 - 新規_総合_国内銀行_vol_6m_lag4 - 正社員数_小売業_大規模_4期前増減率_lag4	- 短期_生産設備_中小企業_小売_差分_lag4 - 貸付残高_件数_4期前からの変動_lag4 - 長期借入金合計金額_小売業_大規模_lag3
建設業	2期先	- 正社員数_建設業_中小規模_lag2 - 短期_生産設備_中小企業_建設_lag4 - 短観_販売仕入_中小企業_建設_差分_lag4	- 長期借入金合計金額_建設業_中小規模_vol_6m_lag3 - 長期借入金合計金額_建設業_中小規模_vol_6m_lag4
製造業	2期先	- 政府支出_差分_lag4 - 日経平均株価_終値_4期前からの変動_lag4 - 貸付残高_件数_4期前からの変動_lag2	- 長期借入金合計金額_製造業_大規模_vol_6m_lag3 - 非正規社員数_製造業_中小規模_lag2

[1] フォワードテストの結果をP.21 参考7-2に記載

Appendix

[参考1] inputデータの詳細①

使用したデータ	出典	URL
TDB倒産件数データ	帝国データバンク	
信用調査報告書	帝国データバンク	
日経平均株価	日経平均プロファイルダウンロードセンター	https://indexes.nikkei.co.jp/nkave/index?type=download
日本政策金融公庫	日本政策金融公庫 融資実績	https://www.jfc.go.jp/n/company/national/g_gaikyo.html
GDP統計政府支出	内閣府 四半期別GDP速報	https://www.esri.cao.go.jp/jp/sna/data/data_list/sokuhou/files/files_sokuhou.html
貿易統計	財務省貿易統計 輸出入額の推移	https://www.customs.go.jp/toukei/suii/html/time.htm
失業率	総務省統計局	https://www.stat.go.jp/data/roudou/sokuhou/tsuki/index.html
民泊住宅宿泊事業届出件数	民泊制度ポータルサイト	https://www.mlit.go.jp/kankochou/minpaku/

[参考2] inputデータの詳細②

使用したデータ	出典	URL
実質実効為替レート	日本銀行時系列統計データ 検索サイト	https://www.stat-search.boj.or.jp/
貸出約定平均金利 (新規・総合・国内銀行)		
貸出約定平均金利 (新規・長期・国内銀行)		
業況DI (中小企業・業種ごと)		
販売価格DI (中小企業・業種ごと)		
仕入価格DI (中小企業・業種ごと)		
雇用人員DI (中小企業・業種ごと)		
設備判断DI (中小企業・業種ごと)		
資金繰りDI (中小企業・業種ごと)		
貸出態度DI (中小企業・業種ごと)		

[参考3-1] 貿易統計の主成分分析（学習データ：～2024Q4）

- 貿易統計の主要48品目のうち、数量・金額の両方が利用可能な27品目を対象とし、主成分分析を用いて単価と数量についてそれぞれ3つの主成分（PC1～PC3）を抽出
- 各主成分の解釈結果から、以下のような意味づけが妥当であると考えられる
- フォワードテストを実施するため、フォワードテスト前までの期間を学習データとして使用

単価の主成分分析の結果

	PC1	PC2	PC3	PC4	PC5
標準偏差	4.54	1.75	1.11	0.93	0.78
寄与率	0.74	0.11	0.04	0.03	0.02
累積寄与率	0.74	0.85	0.89	0.92	0.94

単価_PC1 = 全体の動き
単価_PC2 = 石油関連
単価_PC3 = 部品関連(機械・電子関連)

数量の主成分分析の結果

	PC1	PC2	PC3	PC4	PC5
標準偏差	3.12	2.42	1.73	1.33	1.21
寄与率	0.35	0.21	0.11	0.06	0.05
累積寄与率	0.35	0.56	0.67	0.73	0.78

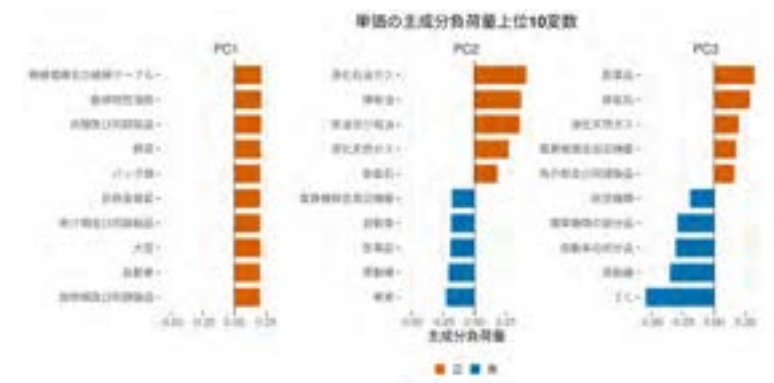
数量_PC1 = 全体の動き
数量_PC2 = 資源関連
数量_PC3 = 食品関連

[参考3-2] 貿易統計の主成分の解釈

- PC1～PC3の各主成分負荷量上位10指標から以下のような意味づけが妥当であると考えられる

単価指標PC1～PC3

- 単価_PC1：異なる品目が同じ正方向に高い負荷量を持つ
⇒ **全体の動き**
- 単価_PC2：石油・鉄鉱石などが正方向に高い負荷量を持つ
⇒ **石油関連**
- 単価_PC3：負の負荷量の航空機、電算機類の部分品関連が強い
⇒ **部品関連**



数量指標PC1～PC3

- 数量_PC1：異なる品目が高い負荷量を持つ
⇒ **全体の動き**
- 数量_PC2：特に資源・重工業関連などが正方向に負荷量を持つ
⇒ **資源関連**
- 数量_PC3：特に食料・素材関連が強い負荷量を持つ
⇒ **食品関連**



[参考4] 過去の予測モデルの変更

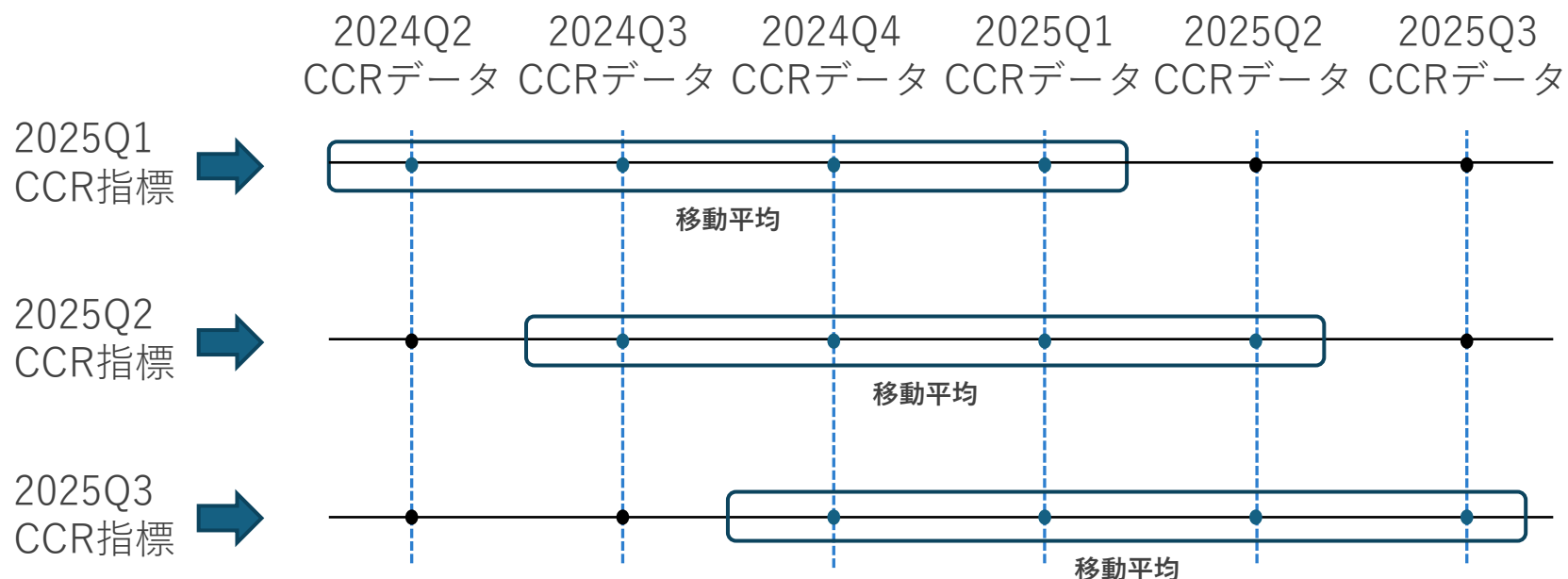
昨今の倒産予測精度の悪化を背景にモデル改良を行った（2025年Q1から運用）

相違点	理論モデル（～2024Q4）	逐次変数選択モデル（2025Q1～）
目的	経済理論に基づき解釈性を重視した予測	機械的な変数選択により直近の状況を強く反映した高精度な予測
変数選択手法	経済理論に基づき説明変数を事前に選定	Stepwise法（変数増減法）により、四半期ごとに変数を選択
変数選択基準	経済理論との整合性を重視して採用（係数の符号が解釈可能か）	コロナ期間を対象としたクロスバリデーション（leave-one-out交差検証法）で平均二乗誤差が最小となる変数の組合せを採用
その他条件	モデルに不均一分散が認められる場合はGLSで推定	変数選択を行う上で、多重共線性、不均一分散、自己相関が認められるモデルは候補から除外
モデルの特徴	経済理論に整合的で解釈性が高いモデル	コロナ禍の影響を強く反映し、四半期ごとの機械的な変数選択によるモデル更新

- 新モデルへの変更により、平均的な予測精度は向上
- しかし四半期ごとに選択変数が大きく変化し、変数の解釈や予測結果の安定性に課題

[参考5] CCRデータについて

- 帝国データバンク社が保有する企業の信用調査報告書^[1]
- 個社ごとに調査日が異なり、大半の企業は1年に一度調査が行われるため、一年間(3期前～当期まで)の移動平均を指標とする
- 長期借入金総額・短期借入金総額・正社員数・非正規社員数を業種ごとに集計



[1]信用調査報告書の詳細. <https://www.tdb.co.jp/services/lineup/research/>

[参考6] Stability Selection^[1]について

学習時点が少ない中で安定した予測を行うため、Stability Selectionを採用

- 一般的な正則化手法による変数選択

全データでLASSO等のモデルで変数選択

0でない係数を持つ変数を採用

<デメリット>

- サンプルやノイズの影響によって選択結果が変わりやすい
- 変数選択結果が不安定・再現性が低い

- Stability Selectionを用いた変数選択

複数回（100～200回）実行

全データの半分をランダムに抽出

各データでLasso等のモデルで変数選択

各変数の選択頻度を集計

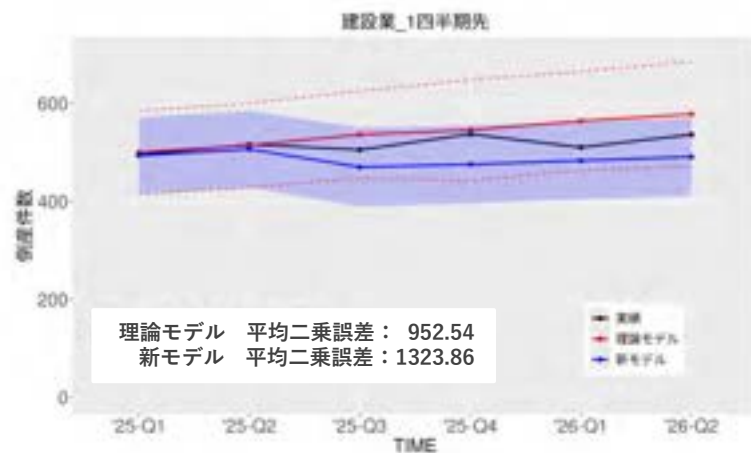
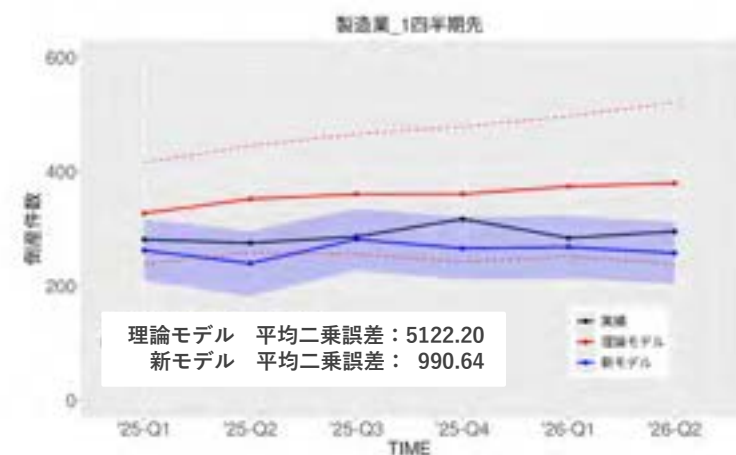
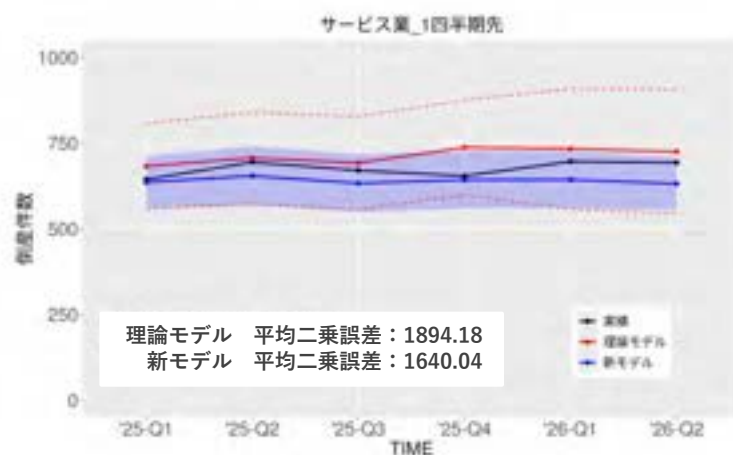
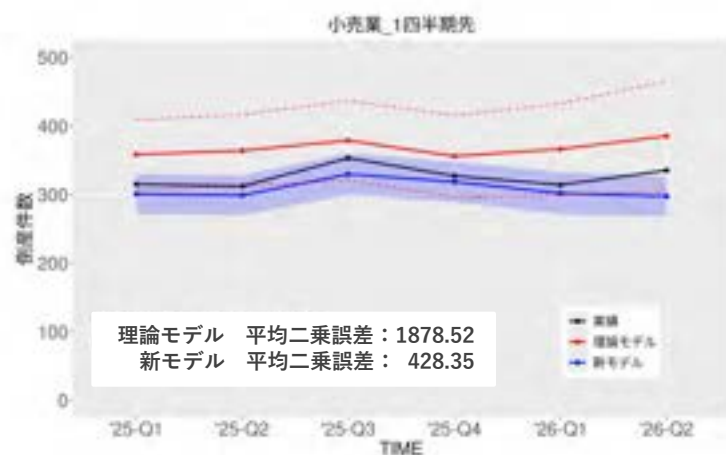
選択確率から採用変数を決定

- 一般的な正則化手法と比べ、データの偶然に左右されて選択される変数を除き、安定して選択される変数を抽出することが可能

[1] Nicolai Meinshausen, Peter Bühlmann, Stability Selection, Journal of the Royal Statistical Society Series B: Statistical Methodology

[参考7-1] 1四半期先モデルのフォワードテスト

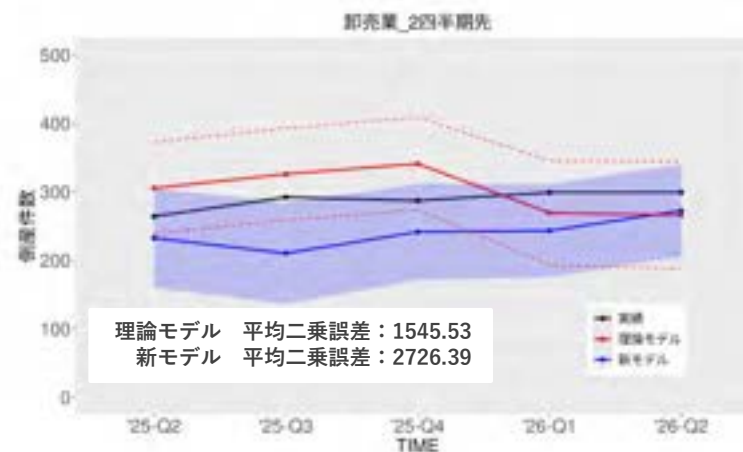
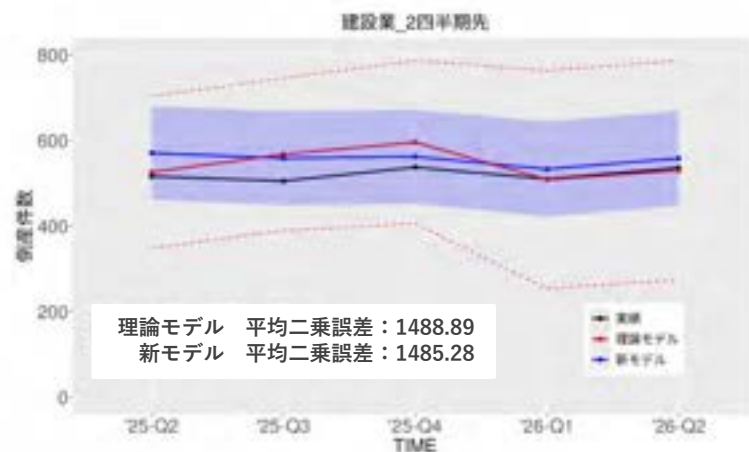
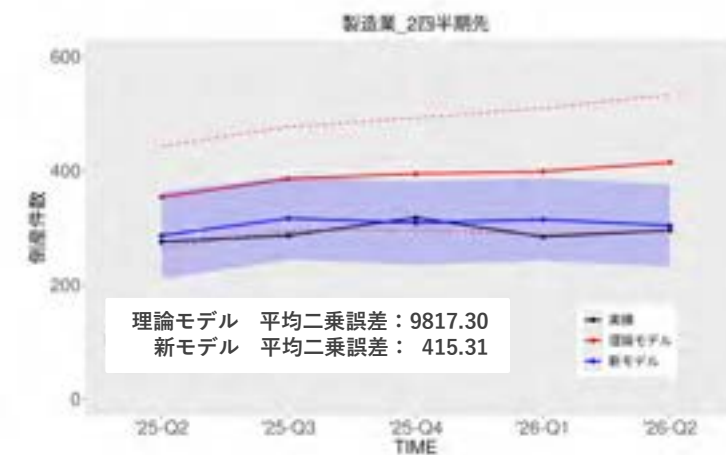
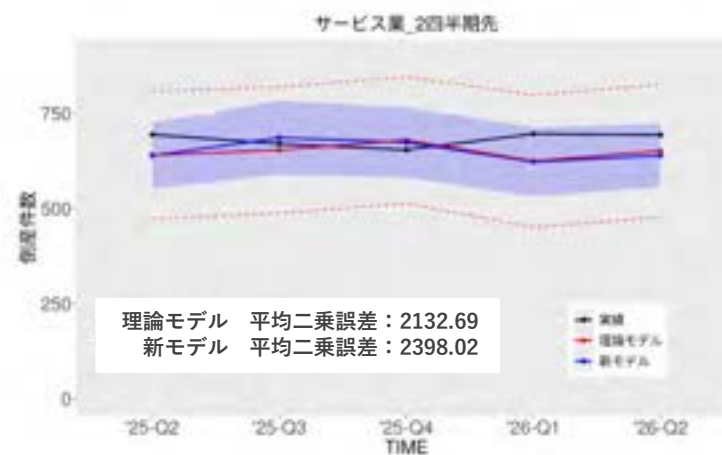
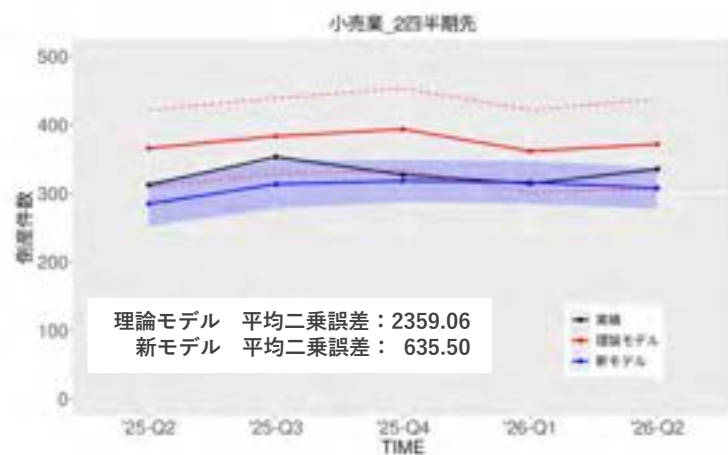
選択された変数を用いて、1四半期先モデルの予測性能をフォワードテストにより検証



- 複数業種で理論モデルを上回る予測精度（MSE）を確認
- 建設業でも予測精度の大幅な低下は見られない

[参考7-2] 2四半期先モデルのフォワードテスト

選択された変数を用いて、2四半期先モデルの予測性能をフォワードテストにより検証



- 複数業種で理論モデルを上回る予測精度（MSE）を確認
- サービス業でも予測精度の大幅な低下は見られない
- 卸売業の予測精度は大きく低下しているため注視する必要がある