

与信モデルのハイブリットAI化

Hybrid AI for Credit Models

ささやきをカタチに。



COMTURE
LEAD THE FUTURE

小野 潔

コムチュア 株式会社

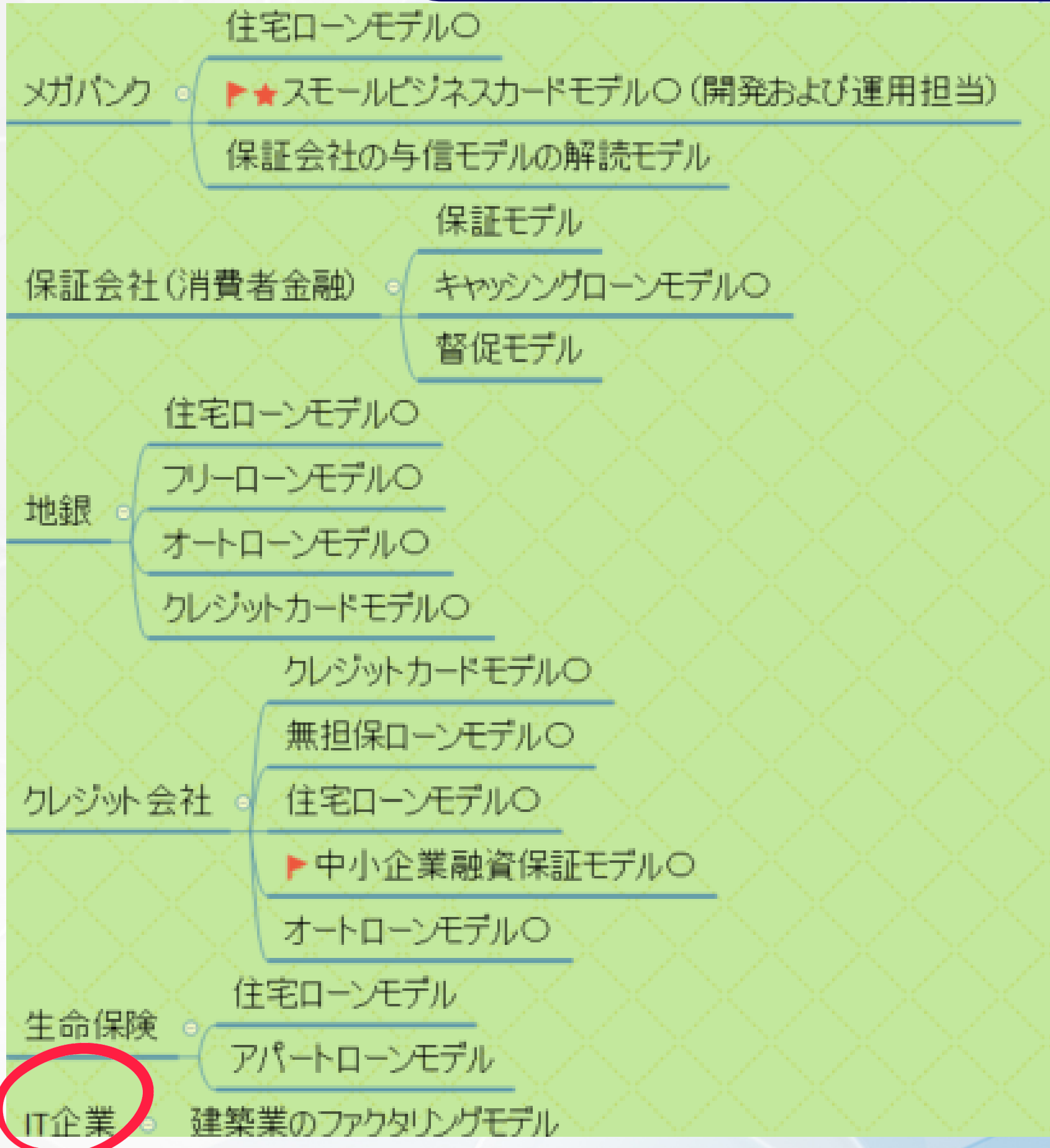
デジタルイノベーション本部

本発表の目的

与信ビジネスの発展において
与信モデルのハイブリッドAI化の
重要性を過去から直近の事例を通じて
説明します。
そして最新事例として、
ファクタリングモデルの構築法を紹介
します。



関与した20の与信モデルプロジェクト



直近は昨年、POC
ファクタリングモデル

1. 与信モデルとは

- 個人または法人の信用力と返済能力を予測
- 融資の適格性や金利を決定



3. 与信モデルのタイプ

- 個人向け: 信用情報や個人属性に基づく精度が高いモデル
- 中小企業向け: 内部の倒産データが不足
- ⇒ 外部企業情報や異なる分類器の組み合わせによる複合モデル



2. 初期与信と途上与信

- 初期与信: 新規融資対象者の信用力を評価
- 途上与信: 追加融資時の信用度の変動を評価
- ⇒ 評価戦略の改善が必要



4. 与信モデルのハイブリッドAI化

- 異種の機械学習、ベイズ統計学、深層学習、自然言語処理、オルタナティブデータ、データベース、クラウド等を組み合わせ
- 高度システム処理や分析を行うAI与信システム



与信モデルの発展史

1990年初頭

米国で機械学習による与信モデルが登場

1997年頃

日本でも銀行で導入検討

1999年

マイニングツールによる個人向け与信モデル開発が始まる

ロジステック回帰、決定木

2003年

中小企業の与信モデル開発が始まる

SVM、アンサンブル学習

2006年

新銀行東京で中小企業融資1000億円累積赤字判明

中小企業モデルの運用失敗

2010年代

個人向け与信モデルを組み込んだ金融プラットフォームが地銀に広まる

不均衡データ理論の体系化
勾配ブースティング

2018年頃

完全自動機械学習アプリによる個人向け与信モデルが登場

XAI (Explainable AI)
オルタナティブデータの活用 (SNS,FB等)

2020年頃

建築業のファクタリングモデル開発が始まる

複合モデルによるハイブリッドAI化

個人向け与信モデルのハイブリッドAI化

同時チューニング

- 相互補完性の課題
 - 銀行と消費者金融は逆戦略
 - 初期与信と途上与信は相互補完的
 - それぞれの最適化が運用上最適とは限らない
 -
- データ充実による可能性
 - データレイクの出現により、ハイパーパラメーターや顧客ランクの閾値等の同時に調整が可能
 - 総合的モデルのパフォーマンス向上

マーケティング戦略への応用

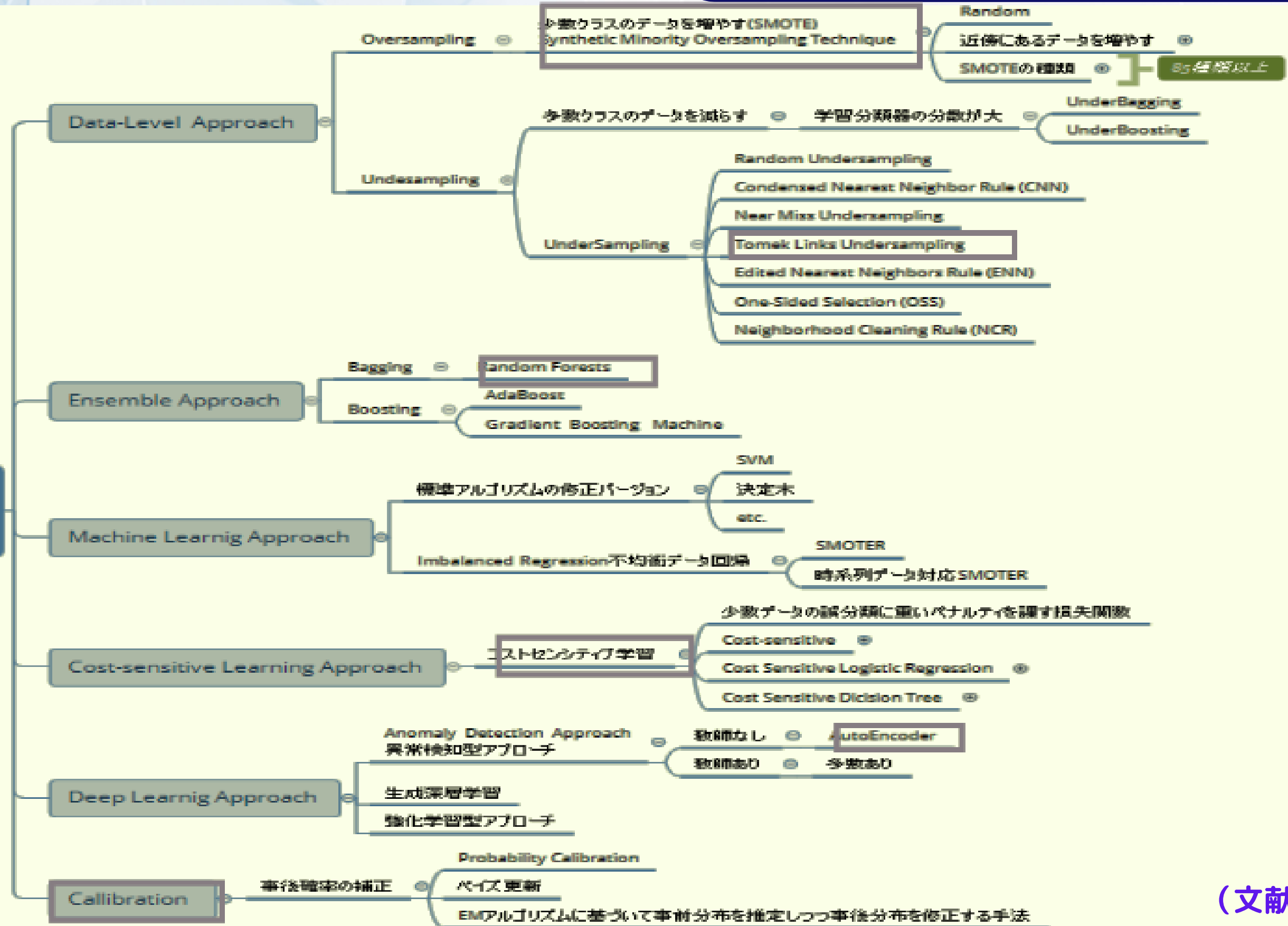
- 半自動チューニングの採用
 - マーケティング戦略も取り入れる
 - 運用現場のフィードバックの反映
 -
- 金融機関の迅速な対応
 - 顧客ランクの閾値、ハイパーパラメータ、平均デフォルト率を外部パラメーターとして設定
 - マーケティング市場変化に迅速対応
 -
- 半自動化の重要性と注意点
 - モデルの適用範囲や限界を理解
 - 運用現場での検証することが重要

不均衡データの 分析理論



不均衡データ分析の理論体系

機械学習の不均衡データの理論体系

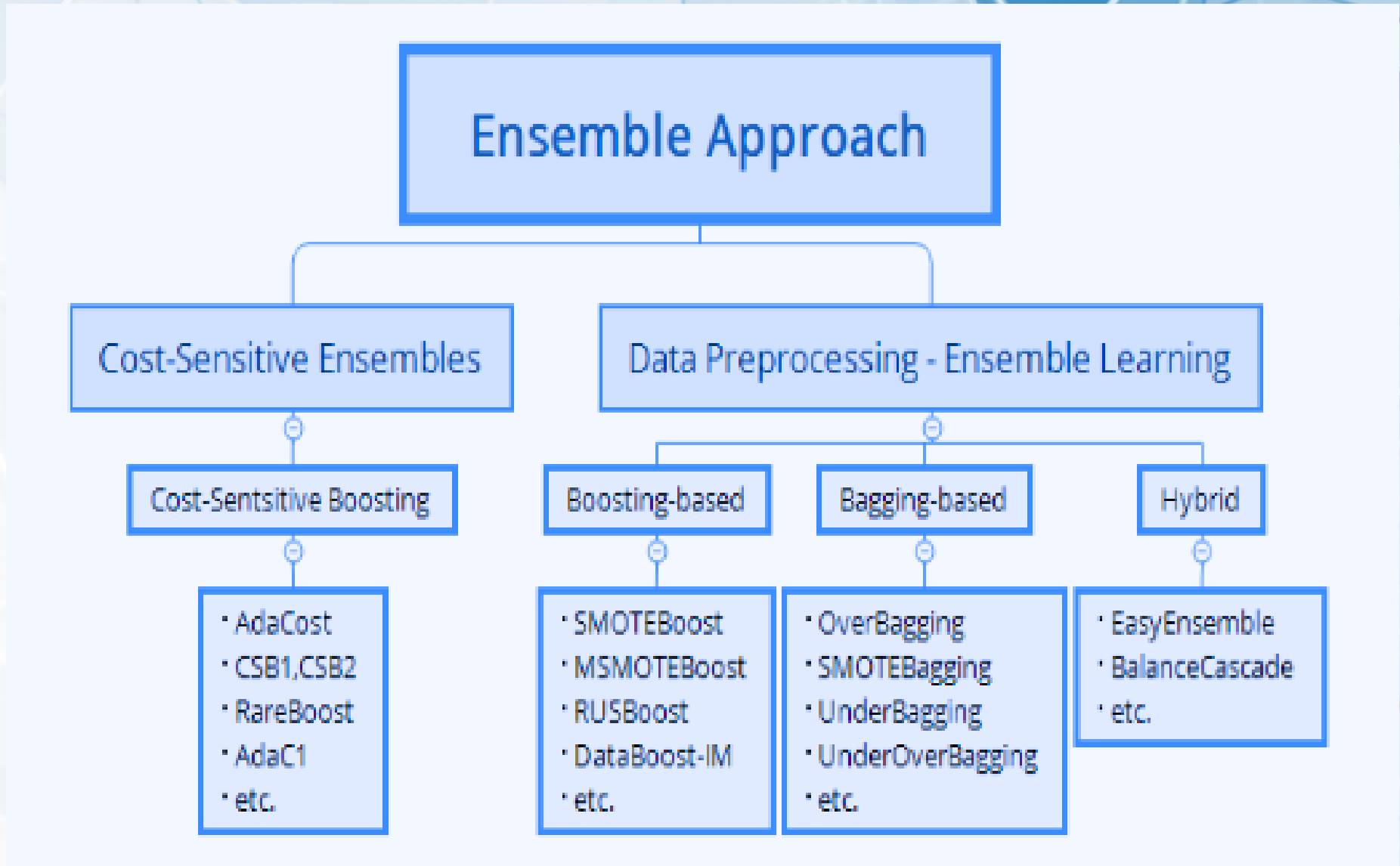
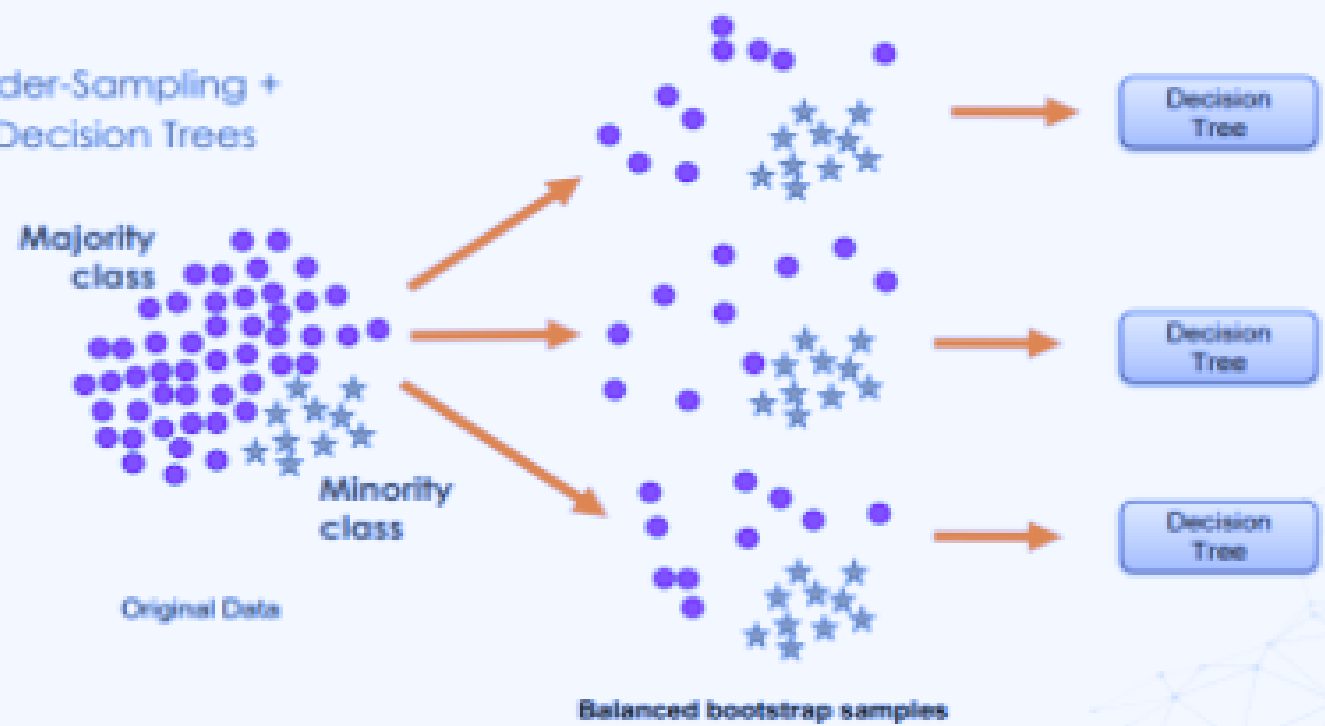


(文献参照)

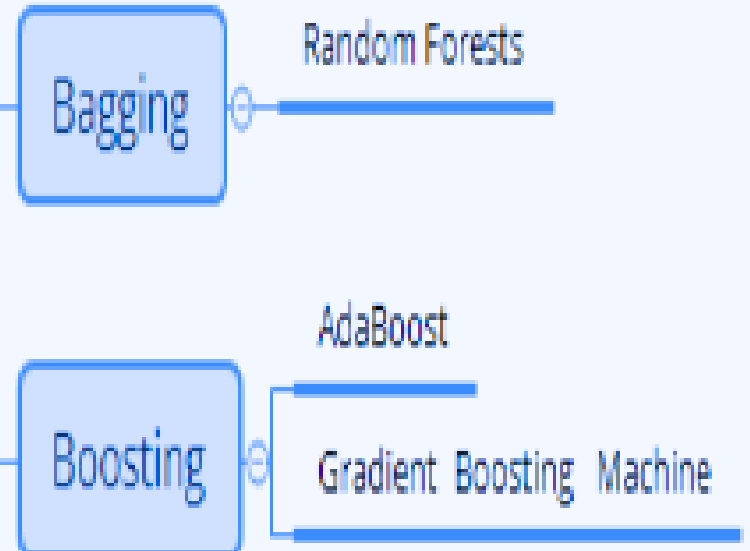
Ensemble Approach

Balanced Random Forests

Random Under-Sampling + Bagging of Decision Trees

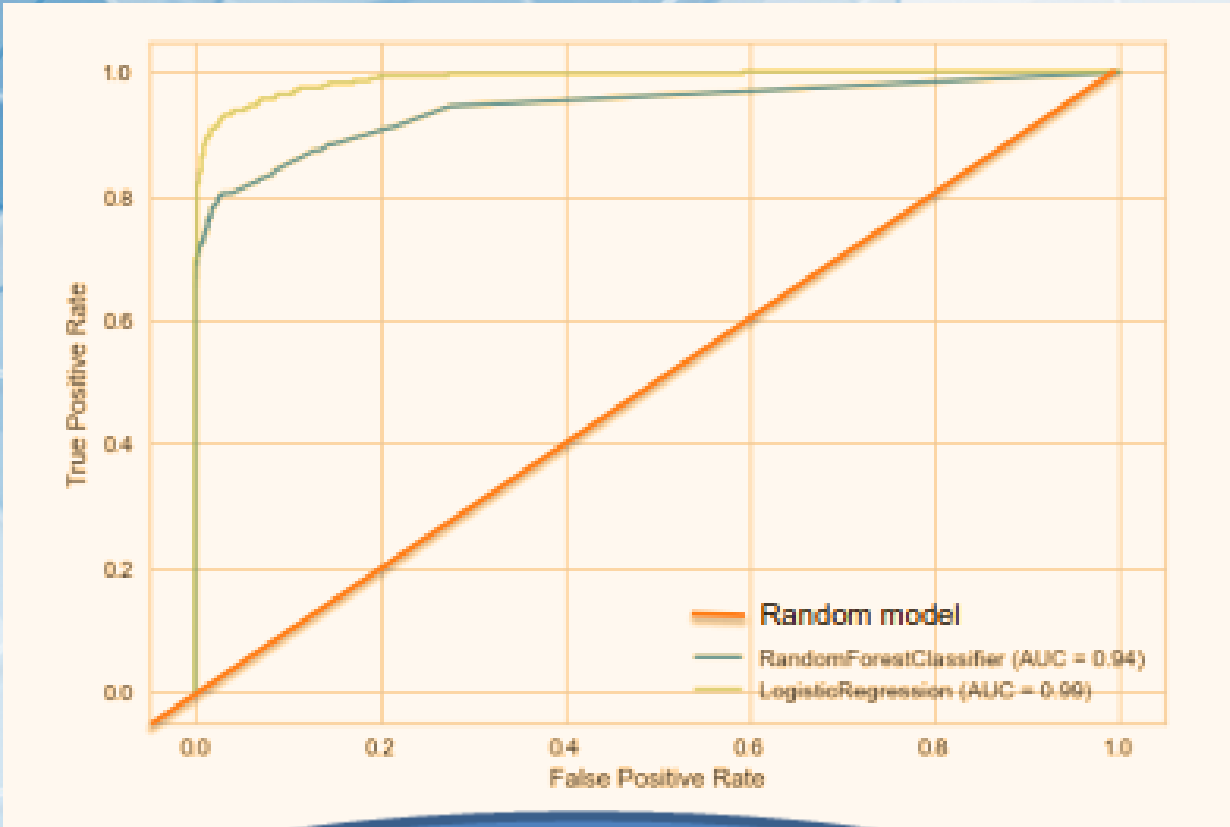


Ensemble Learning

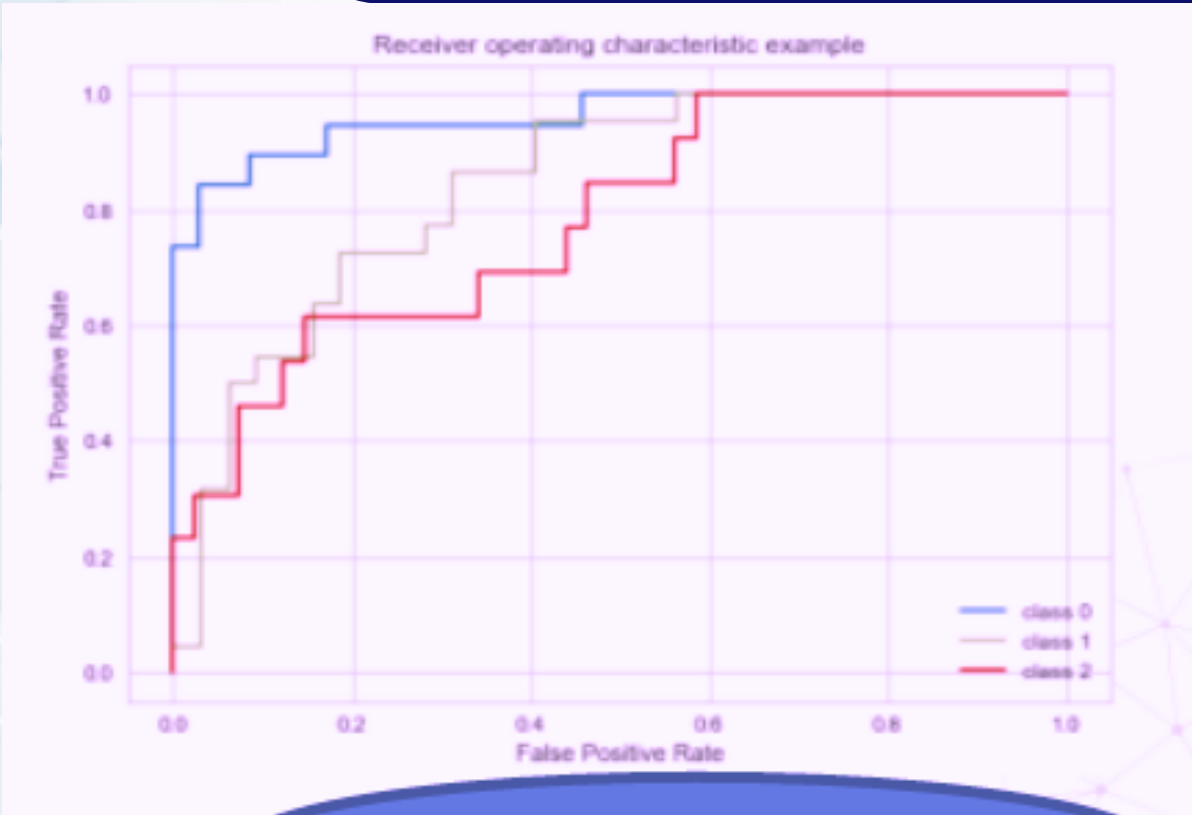


(文献参照)

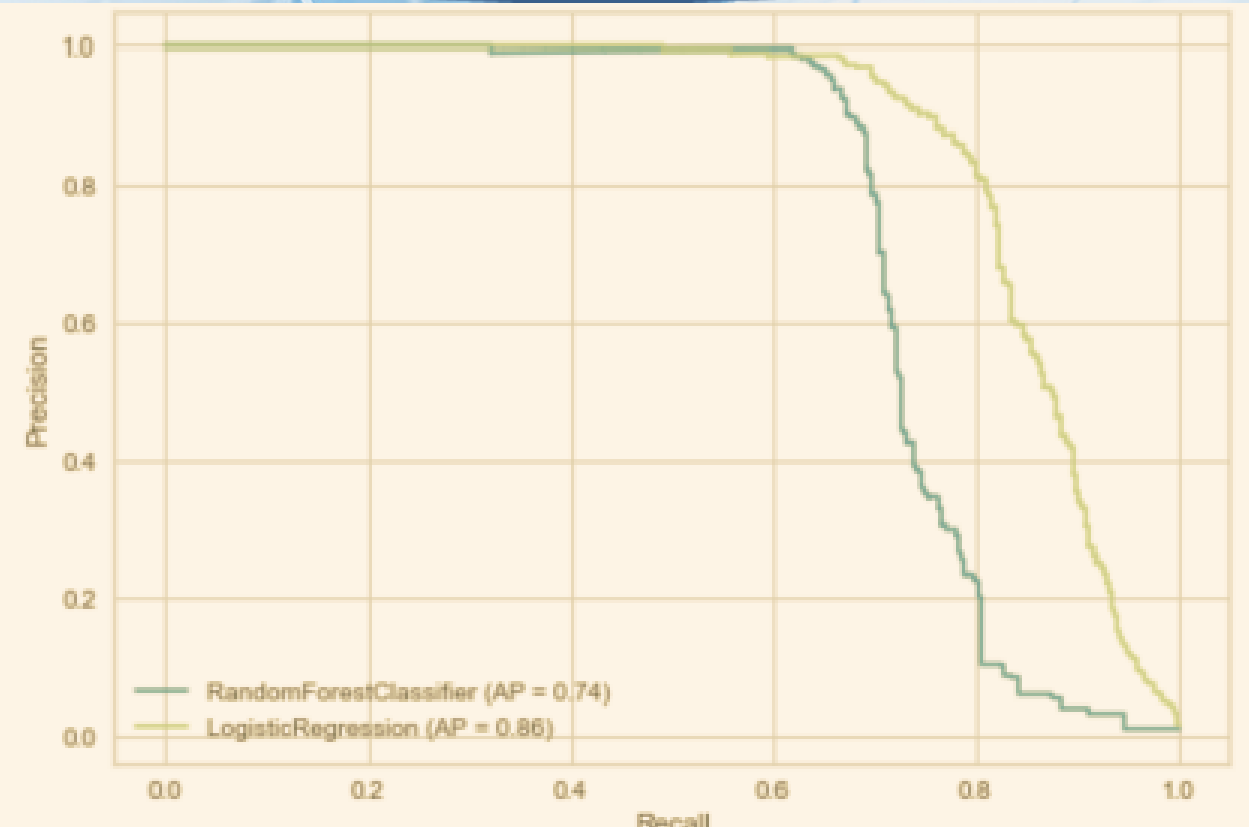
不均衡データのROC & PR曲線



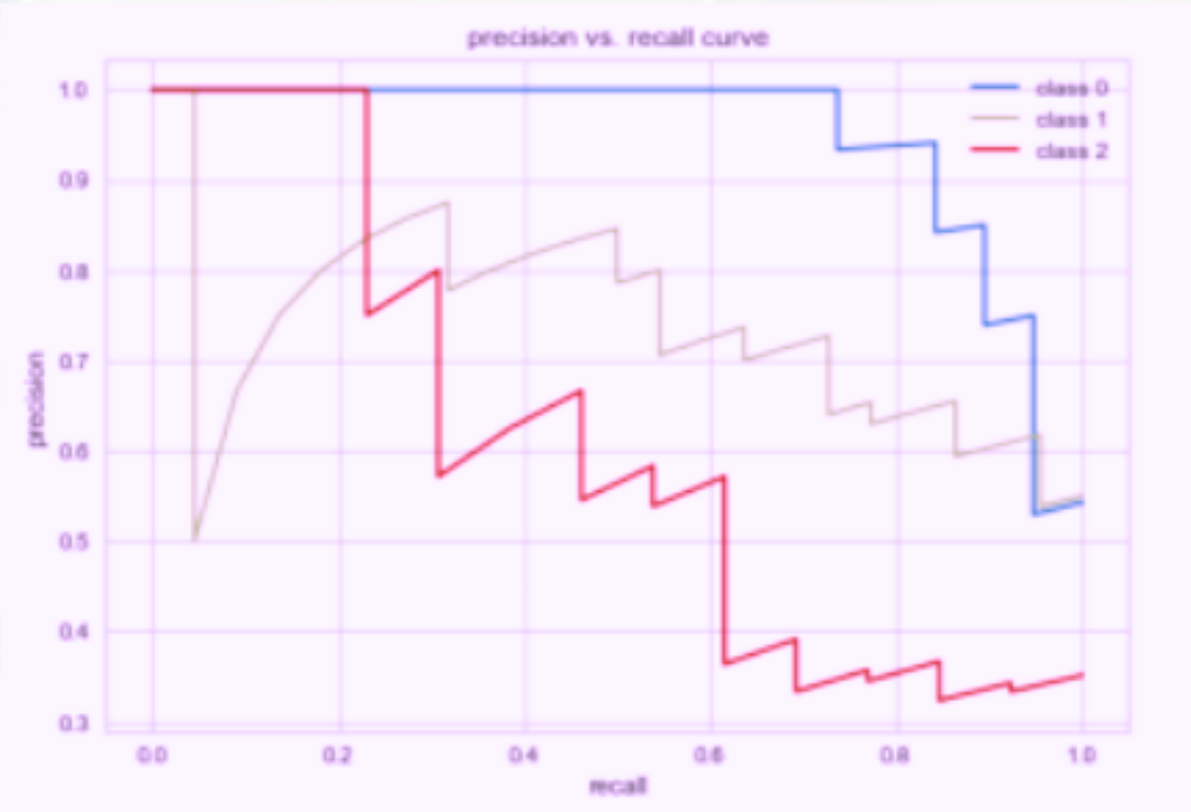
均衡データ



負事例少数データ



PR曲線



(文献参照)

中小企業の与信モデルのハイブリッドAI化

不均衡データの機械学習の発達

- モデル評価指標
F2, Recall 値, PF曲線
- 不均衡データ理論の展開
OverSampling: SMOTE
UnderSampling:
TOMEK LINKS

F-measure

$$F_{\beta} = \frac{(1 + \beta)^2 \cdot \text{recall} \cdot \text{precision}}{\beta^2 \cdot \text{recall} + \text{precision}}$$

機械学習とベイズ統計のハイブリッド AI 化

1. 中小企業モデル

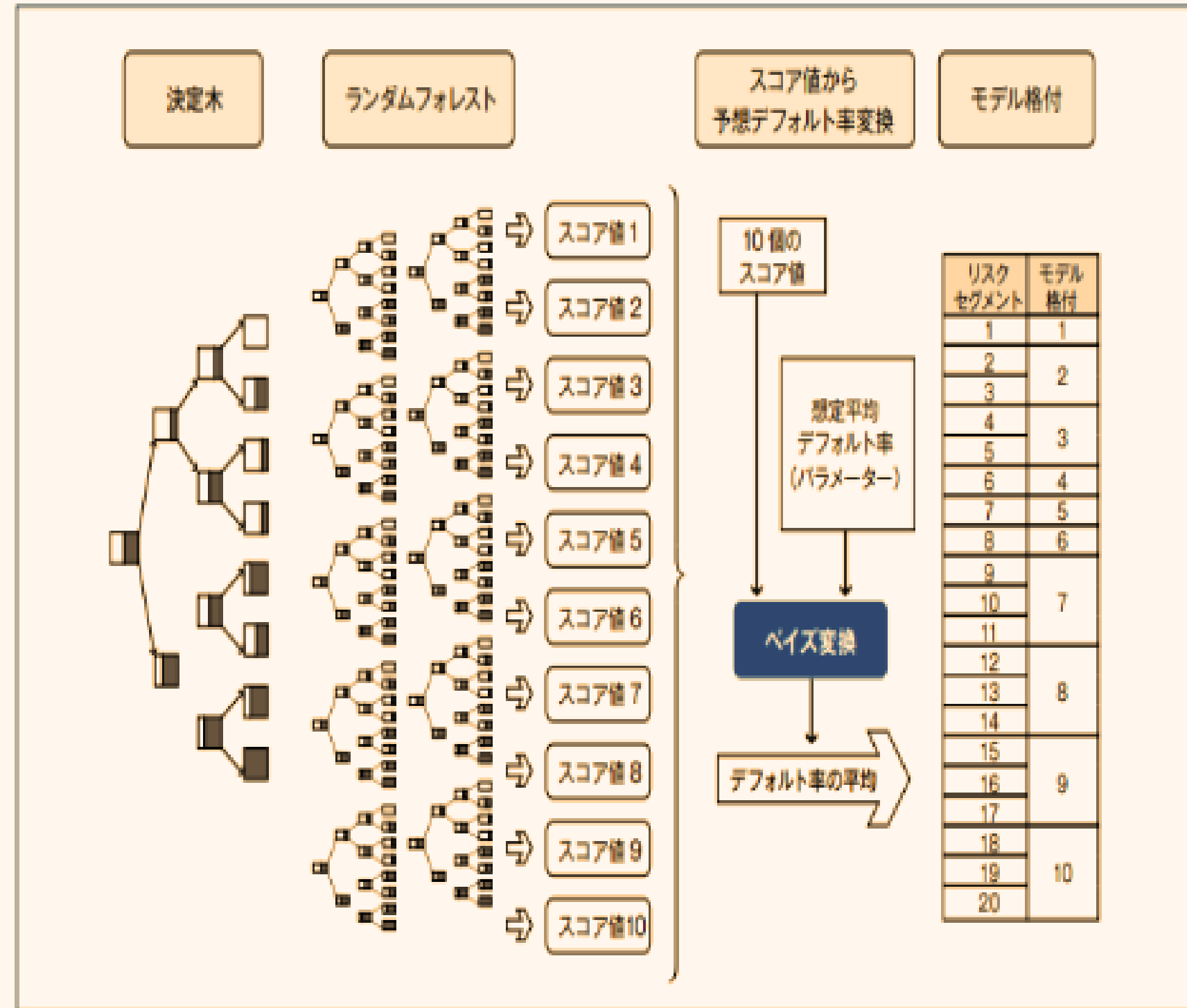
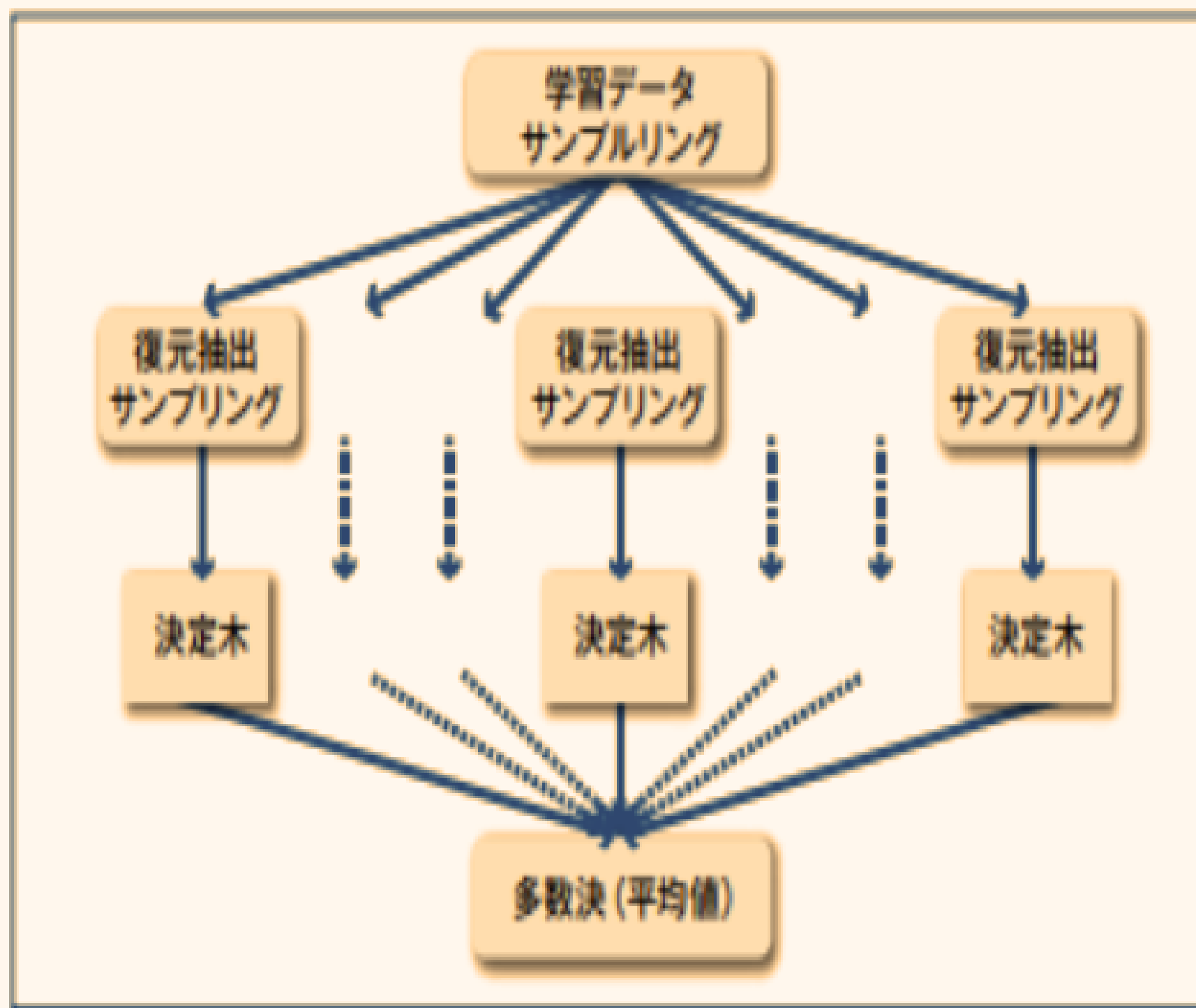
- サンプルング: Under-Sampling (正事例:負事例=1:1)
- 分類器: Random-Forest
- 倒産確率: ベイズ推定
-

2. ベイズ推定

- 母数の平均倒産率と確信度に基づく
- 全国の中小企業倒産率と相関
-

3. 経済状況への対応

- 経済変動をモデルに取り込み個別企業の倒産確率に反映
- モデル再構築なしで応急対応



建築業のファクタリングモデルの課題

複合モデル

- 中小企業モデル
社内倒産情報
外部企業情報(TDB)
- スケジュール遅延モデル
労務安全情報
- 連鎖倒産モデル
外部企業情報
労務安全情報
その他

連鎖倒産モデル

- 系列構造の取り込み
GNN (Graph Neural Network) の検討

ファクタリングモデルの拡張

- オルタナティブデータの活用がKEY
- 商業のファクタリングモデルへの拡張

オルタナティブデータの活用

- 顧客の特性の把握
購買履歴
SNSの投稿履歴
- 信用情報ない移民・
難民対応
- 建築業の労務安全データ
グリーンサイトデータ

ブロックチェーン技術の導入

クレジット申込の途中取りやめは手入力の面倒が最大の課題

- 個人情報の漏洩や改ざん防止

米国POS (Point Of Sale) ファイナンスの拡大

- 米国信用情報機関の新サービス
顧客の申し込み時に個人情報を収集し、自動入力サービスを提供、手入力作業の7割削減
- IT企業の台頭
- 日本の情報機関の対応

自然言語処理の検討

- 営業日誌やSNSの分析
自然言語処理
BERTやGPTの利用



与信モデルのハイブリッドAI化が進んでおり、過去と直近の事例を含めてまとめた。さらに米国の現状を報告し、日本の与信ビジネスの留意点を考察した

今後の与信モデルの開発の指針になれば幸いである。