

※この資料は、東京大学グローバル消費インテリジェンス寄付講座（GCI）2024 Winterの講義で使用したものです。

機械学習を「社会実装」するということ

2025年版

— 思考せよ、創造せよ —

Moe Uchiike

自己紹介 - 内池 もえ -

- (株)ブレインパッド エンタープライズユニット ソリューションリード / シニアマネジャー
エグゼクティブデータサイエンティスト
- PMやモデル開発リーダーとして、複数の機械学習/最適化プロジェクトをローンチまで推進
- 松尾研究室 DL4USの2期生。最終課題のテーマは 「おでんの需要予測」
- GCIでの講義は8度目
- 日本ソムリエ協会認定ワインエキスパート



本日は、**現場の前線に立つ実務家の立場**からお話しさせていただきます！

本講義のねらい

本日の講義を、3つの解像度を上げるきっかけにさせていただきたいと考えています。



なぜやるのか



何に立ち向かうのか



どうやるのか

目次

アイスブレイク	データ活用でどんな価値を提案する？
第 1 章	AIプロジェクトの位置づけ
第 2 章	2025年の機械学習プロジェクト
第 3 章	社会実装を阻む「罨」と、その解決策
第 4 章	AI大活用時代に何ができるか
まとめ	本日お伝えしたかったこと

アイスブレイク

データ活用で
どんな価値を提案する？

アイスブレイク データ活用でどんな価値を提案する？ (1/2)

以下の状況を踏まえ、考えたことをアウトプットしてみてください。

時代背景



- 世はまさに**大生成AI時代**。世界中が開発競争に参戦し、日夜GPUを回している
- GPUが巨大な処理をする際に大量の電気を消費するため、**電力需要は増加見込み**
- 一方で、2050年のCN (カーボンニュートラル) に向けて**CO2削減目標は必達**

皆さんの役割



- 顧客に伴走しデータ活用を提案・支援する企業に勤めるデータサイエンティスト
- データ活用の力で産業をアップデートするのがミッション

現在の状況



- 電力会社A社とデータ活用の方向性についてディスカッションすることになった
- A社との取引の実績はなく、今回のディスカッションが初接触

hint

- CO2を排出せず、枯渇とも無縁な**再生可能エネルギーによる発電**が注目を浴びている (太陽光、風力、地熱、バイオマス 等)
- 電力需要にはピークがあり、上手く**ピークシフト**することで需給バランスの調整を図ることができる
- 卸電力市場や次世代電力網など、需給に合わせて**電力を融通する仕組み**が整備されつつある
- **電力会社の業務は発電以外にも色々ある**

皆さんならA社とのディスカッションの場で、A社に対してどのような取組/価値を提案しますか？

※時代背景を踏まえた提案に限定せず、電力会社をデータ活用で支援する方法を広く考えてほしい

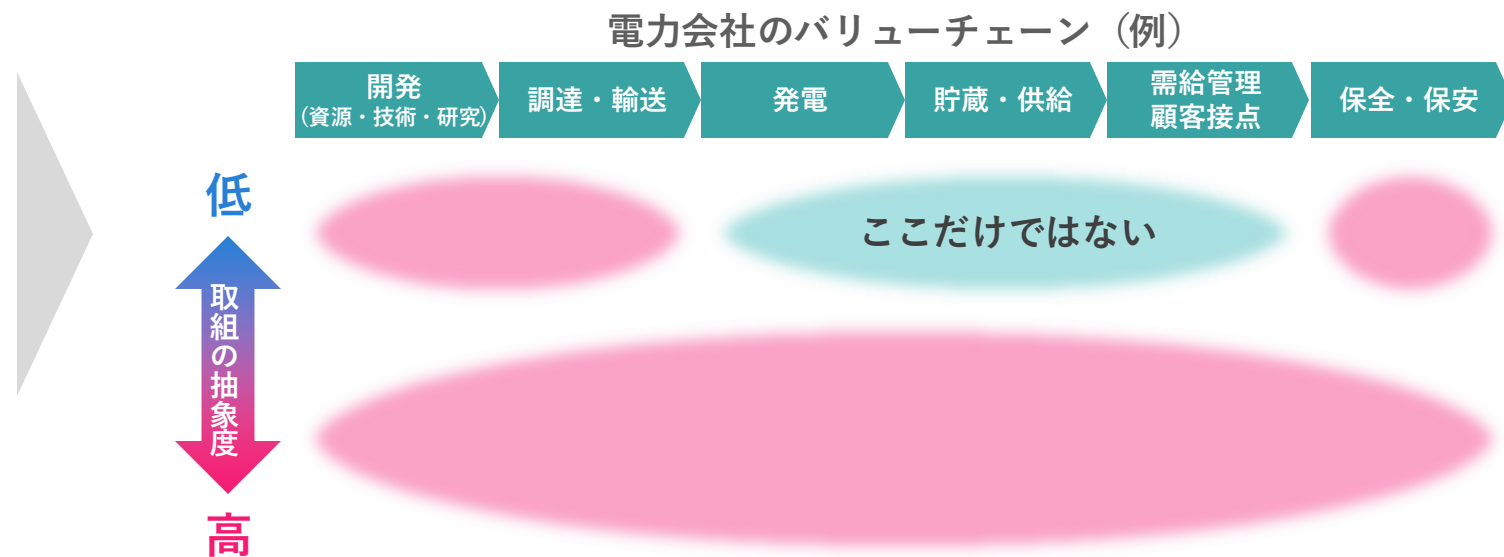
アイスブレイク データ活用でどんな価値を提案する？ (2/2)

この問題に明確な答えがあるわけではありません。
ですが、思ったよりも広い視点でデータ活用の価値を提案する余地があることがわかります。

考えられそうな方向性

- 再生可能エネルギーによる発電方法の多様化をデータ活用で支援する方法はないか
- 電力の需給調整弁をデータ活用で支援する方法はないか
(電力需要予測、蓄電池の運転制御、電力市場との取引最適化)
- 単純にモニタリングができるだけでも価値があるのではないか
(再エネによる発電量 等)

でも、こうすると……？



データ活用で課題を“解く”だけでなく
価値を“創造”することが我々のミッション！

第1章

AIプロジェクトの位置づけ

第1章：AIプロジェクトの位置づけ プロジェクトとは

プロジェクトは、「有期性」と「独自性」の2つの特徴を持ちます。

	有期性	独自性
プロジェクト	あり	あり
通常業務	なし	なし

特定の目標を達成する計画を立て、**限られた時間の中で**成果を出していくのがプロジェクト

プロジェクトについてより深く学びたい方は、PMBOK第7版や関連書籍に目を通してみてください

第1章：AIプロジェクトの位置づけ 様々なAIプロジェクト

AIプロジェクトといっても、その在り方は様々です。本講義では、特に断りがない場合「サービスインをゴールとするBtoBの機械学習プロジェクト」を想定して話を進めます。

	プロジェクトのゴール	ビジネスモデル	要素技術
本講義の 主な話題	サービスイン	BtoB	機械学習
	PoC	BtoC	数理最適化
	示唆出し	CtoC	etc.
	コスト削減	etc.	
	etc.		

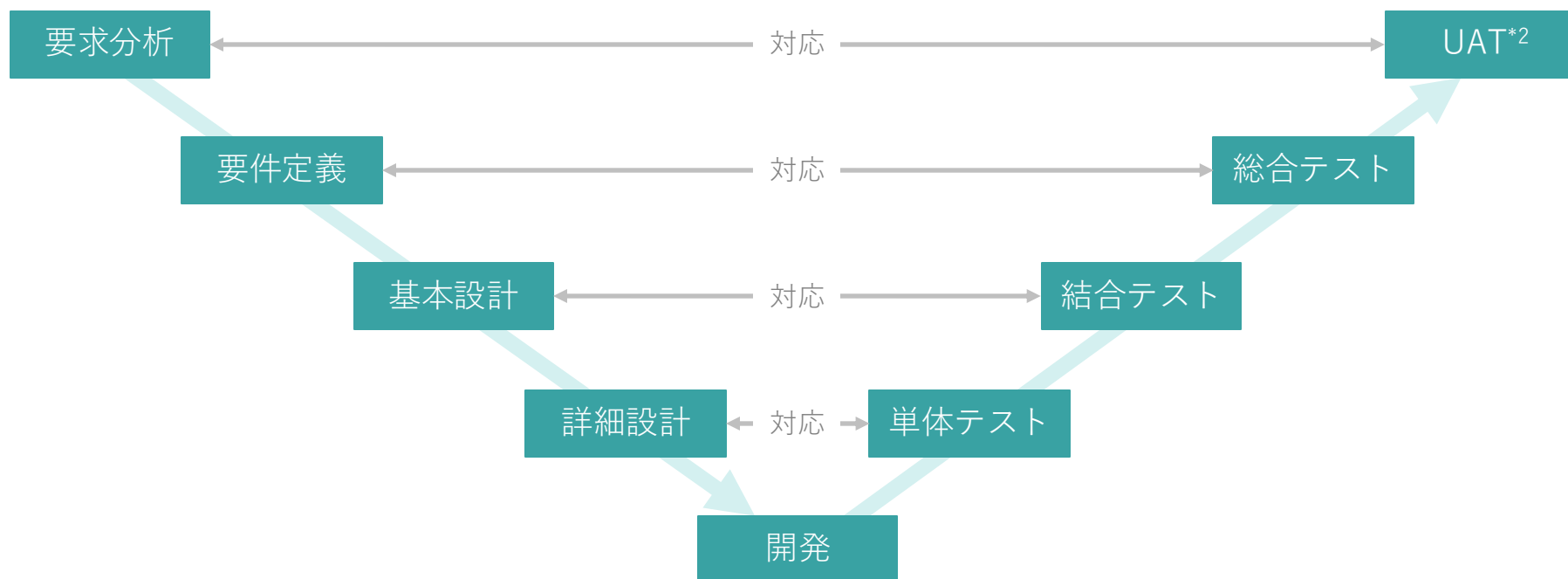
上記に加え、インダストリー（金融業/製造業等）や取り組みテーマ（需要予測/画像認識等）も様々

昨今はLLMなどの生成AIの活用を前提としたプロジェクトも立ち上がりつつある

研究機関の研究とは目的が異なることにも注意が必要（新規性よりも効果が出ることや使い続けられることが重視される）

第1章：AIプロジェクトの位置づけ 一般的なプロジェクトのプロセス

一般的なプロジェクトは、しばしば「V字モデル*1」と呼ばれる開発工程を踏んで進められます。

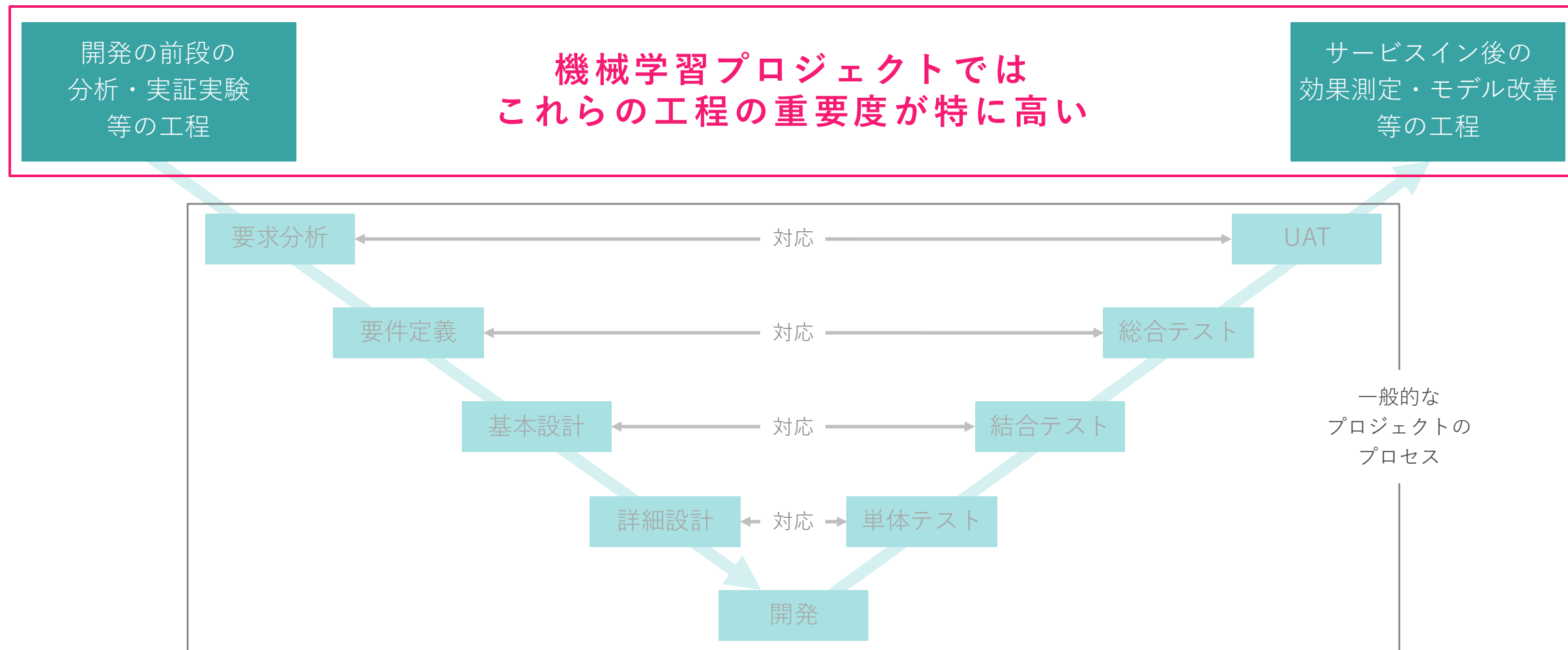


*1: 主にウォーターフォール型開発で用いられる概念

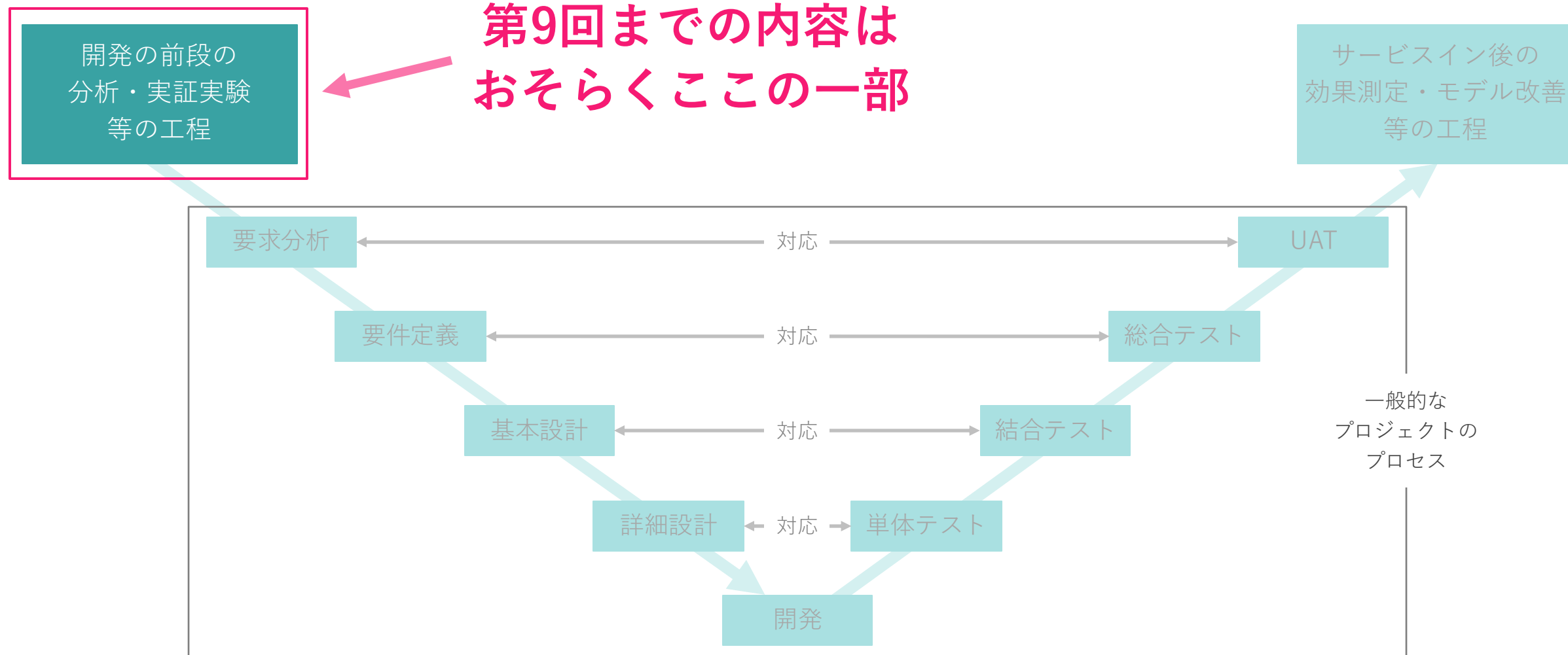
*2: User Acceptance Test。このテストを通過しなければ検収されない

第1章：AIプロジェクトの位置づけ 機械学習プロジェクトのプロセス (1/2)

機械学習などを要素技術とするプロジェクトは、一般的なV字モデルで完結しません。



第1章：AIプロジェクトの位置づけ 機械学習プロジェクトのプロセス (2/2)



ここまで大変なことを

ここまで大変なことを
学問と向き合い、技術を磨いてまで

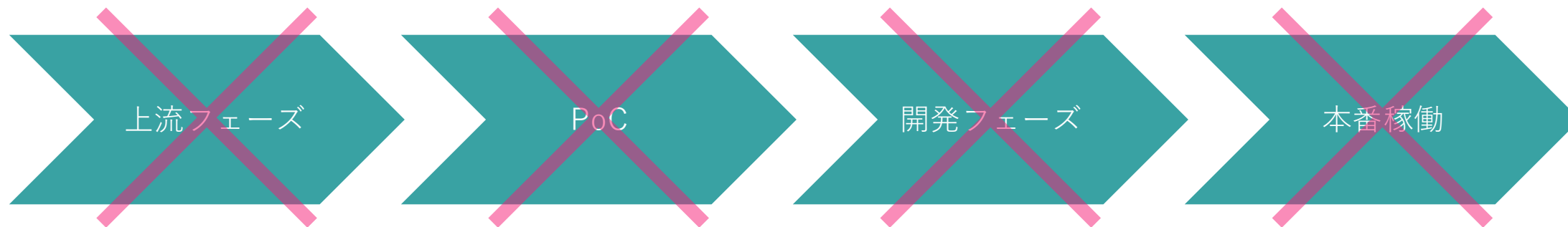
なぜやるのか？

第2章

2025年の 機械学習プロジェクト

第2章：2025年の機械学習プロジェクト ほとんどのプロジェクトが社会実装されずに終わっていた

かつて、研究が進むことと社会で活用されることの間には大きな溝がありました。^{*1}



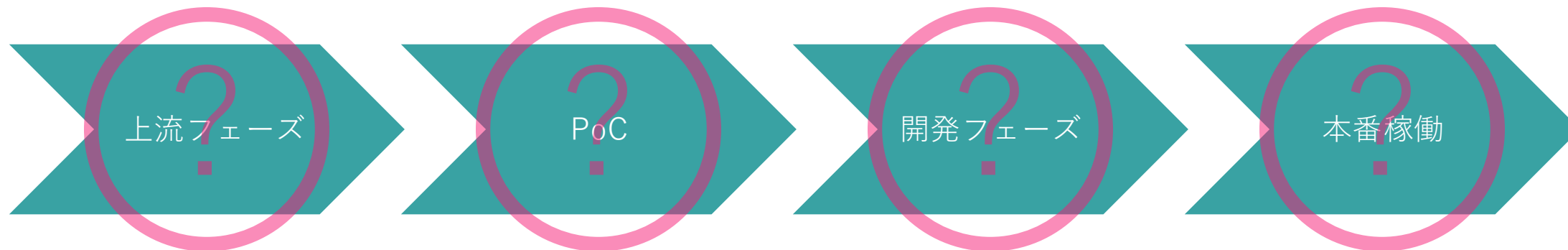
- ✓ アーリーアダプターが実験的にPoC^{*2}に取り組むという側面があった
- ✓ しかし、PoCはそう簡単に上手くいくものではなかった
- ✓ PoCが上手くいったとしても、その後のフェーズも一筋縄ではいかなかった
- ✓ 遡って、「上流フェーズでの問題設定が適切ではなかった」ということも多々あった

*1: 【参考】 Rob van der Meulen and Thomas McCall. *Gartner Says Nearly Half of CIOs Are Planning to Deploy Artificial Intelligence*. Gartner. 2018. <https://www.gartner.com/en/newsroom/press-releases/2018-02-13-gartner-says-nearly-half-of-cios-are-planning-to-deploy-artificial-intelligence>, (参照2025-01-07)

*2: Proof of Conceptの略語で、実証実験(実際に上手くいくか確かめる活動)のこと

第2章：2025年の機械学習プロジェクト ほとんどのプロジェクトが社会実装されずに終わ…らない！？

「社会実装」と呼べる成果を目指して、いかに構想・実行するかが問われる時代へと突入しました*1。
しかし、依然としてそれを阻む「罫」は存在。実用段階ならではの課題に直面する機会も増えてきました。



- ✓ 単に使えるだけでなく、より広い層への浸透や、社会のアップデートを伴う成果が期待されるように
- ✓ データ/モデル活用の方法論が進化、画期的な技術も複数登場*2
- ✓ プロジェクトにおける経験知の蓄積や、プラットフォームの整備が加速
- ✓ 技術・方法論・インフラの進化を以てしても、依然として難度は高い。実用段階ならではの課題も

*1: 既に成功/失敗を経験したアーリーアダプターによる2周目の取り組みや、第2集団による競争劣位からの脱却を意図した取り組みが加速。実験で終われないプロジェクトが増えた

*2: 2022年はStable Diffusion、Whisper、ChatGPTなどの登場に限界が沸き、2023年以降はLLMそのものや、周辺技術と組み合わせたサービス開発の競争が激化。ついに一般層にも普及した

第2章：2025年の機械学習プロジェクト 業界トレンドと求められるスキル

【凡例】

業界トレンド

✓ 求められるスキル

2025年現在、ビジネスの現場における機械学習には「社会実装」が求められるようになりました。土台となる知識（統計学や機械学習）を獲得し、変わりゆく業界トレンドを踏まえてスキルを磨いていくことが必要です。



システム化やDXの流れで、プロジェクトは大規模化・複雑化

- ✓ 3つの領域（ビジネス力、データサイエンス力、データエンジニアリング力）をまたいだ経験知
- ✓ より高いレベルの構想力、実行力



単に使えること以上の成果への期待の高まり

- ✓ UI/UX等の人間との接点を洗練させるスキル（あるいは必要な人材を迎え入れた上でのチームワーク力）
- ✓ プロジェクト終了のその先や波及効果までを見据えて取り組み全体をデザインする力



第2集団によるデータ利活用の本格化

- ✓ 組織の文化や状況、ケイパビリティを踏まえて施策をデザインする力
- ✓ 抽象的/不完全な情報を紐解き、具体的な形に落とし込んでいく力



ツールや手法、方法論の急速な進化と多様化

- ✓ ツールや手法、方法論の膨大な組合せから課題に適したアプローチを選び抜き、組み合わせて応用する力
- ✓ 常識にとらわれず次々とアイデアを生み出す思考の柔軟性



いわゆる「AI倫理」に対する社会的要請の高まり

- ✓ 倫理的な観点で問題を捉え、現実的な解を導くための知識と思考力
- ✓ サービスを形にする際に起こり得ることを事前に想定するための経験の幅広さと想像力



競争の激化により、AIというだけでは勝てない時代に

- ✓ 批判的思考をもって、自ら問いを立て、言語化する力
- ✓ 独自の付加価値をつけていくための発想力や、多様性の確保などによって発揮されるユニークスキル

第2章：2025年の機械学習プロジェクト データサイエンスの現場最前線では

構想策定ニーズの高まり

- データサイエンスは個別の解決策ではなくDX構想の要
- 失敗経験のある企業の「2周目需要」も盛況
- 新たな座組や方法論など「名前のない仕事」の要請

データ人材ニーズの変化

- 構想を描ける人材、問いを立てられる人材に熱い視線
- コードが書けるだけでは代替可能
- データの「使い手」が求められている

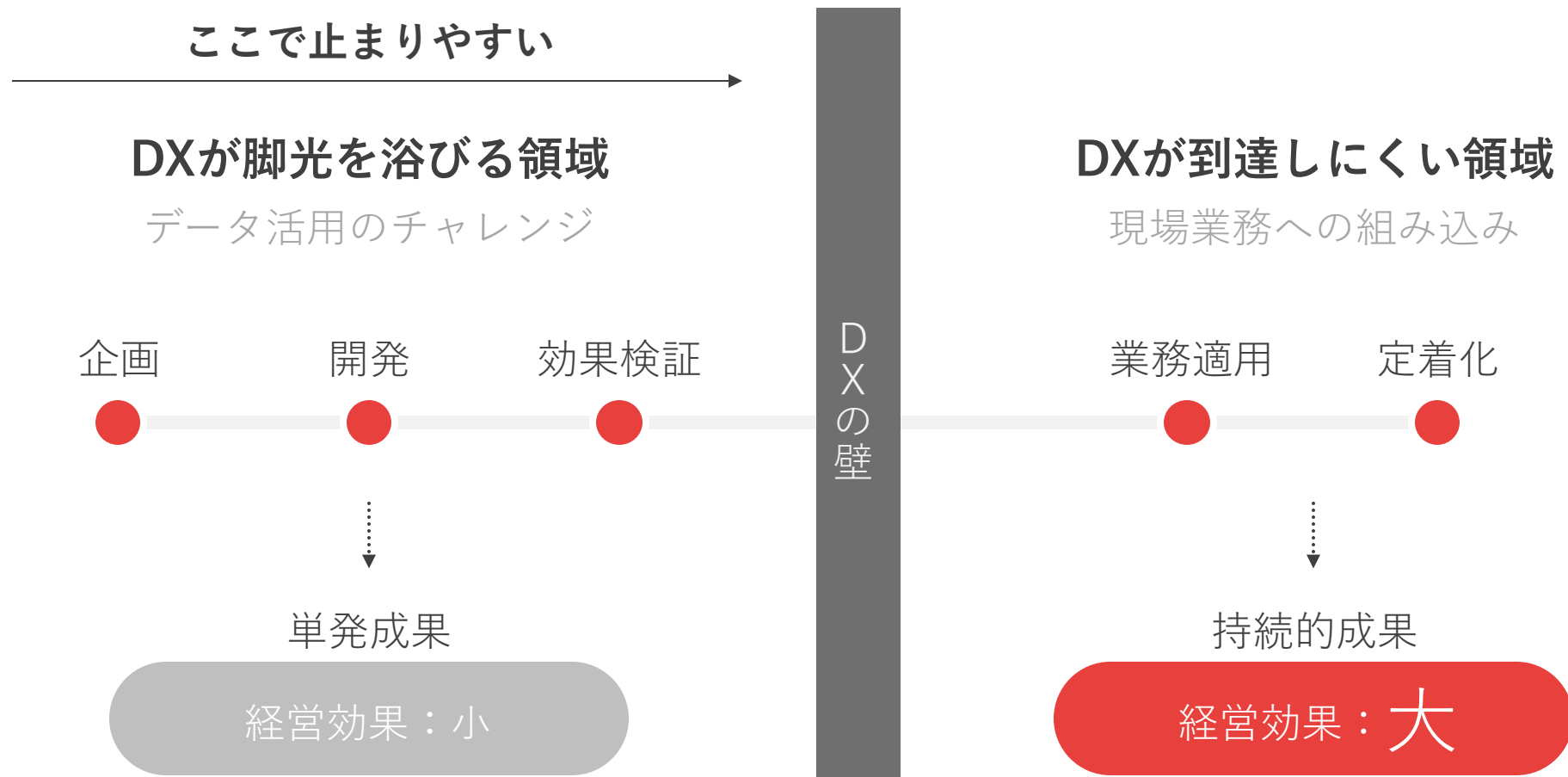
- AI/データサイエンスの応用範囲・応用可能性が急拡大
- 特定分野の専門性を掛け合わせた事例が次々に登場

Data Science × X

- 大企業を中心にデータ活用はさらに加速
- ソリューションの民主化・次世代標準化は道半ば
- AIディバイドが徐々に顕在化

データ活用格差

第2章：2025年の機械学習プロジェクト 【参考】DXの壁



生成 AI 時代

第2章：2025年の機械学習プロジェクト LLMをはじめとする生成AIの概況

LLM（大規模言語モデル）をはじめとする生成AIの急速な進歩により、AIを取り巻く環境は一変しました。

一般層への普及	非専門家でもChatGPTぐらいは使ったことがある or 聞いたことがある状況
	学生のLLM利用に対する立場を表明する教育機関/教員が出現
日本中・世界中が注目	2023ユーキャン新語・流行語大賞TOP10に「生成AI」が選出
	OpenAI CEO サム・アルトマン氏の発言や進退が世界中でニュースに
既存のタスクの代替は 既に一部実現	働き方の常識が徐々に変化 <ul style="list-style-type: none">例：資料の目次作り、議論の観点出し、壁打ち相手として活用例：GitHub Copilotを利用して効率的にコーディング
できることが多様化	GPT-4等のLLMに加え、生成AIをベースにしたクリエイティブ用途のサービスも増加
	APIやプラグインによる機能拡張が一般的に
	「独自にカスタマイズしたGPT」をノーコードで開発することが可能に (ChatGPTのGPTs)
開発競争の 加速・熾烈化	OpenAIのChatGPTは、画像解析機能を備えたマルチモーダル版モデルやGPTsのリリースでより便利に
	AnthropicはClaude2を、GoogleはGeminiをリリース
	国内ではMetaのLlama 2等をベースにした日本語強化モデルの模索が進捗
	LLMの性能評価手法の検証や、新たな手法の提案が進捗
AIガバナンスの 重要性の再発見	LLM等の生成AIの普及に伴い、実用段階ならではの問題が明らかに <ul style="list-style-type: none">例：ハルシネーションや倫理的な問題、プロンプトインジェクション攻撃などの新たな脅威例：なりすましの問題、プライバシーの保護、権利の問題や著作者の心情への配慮

第2章：2025年の機械学習プロジェクト LLMをはじめとする生成AIの概況

LLM（大規模言語モデル）をはじめとする生成AIの急速な進歩により、AIを取り巻く環境は一変しました。

一般層への普及	非専門家でもChatGPTぐらいは使ったことがある or 聞いたことがある状況
日本中・世界中が注目	学生のLLM利用に対する立場を表明する教育機関/教員が出現
既存のタスクの代替は既に一部実現	2023ユーキャン新語・流行語大賞TOP10に「生成AI」が選出 OpenAI CEOサム・オルトマン氏の発言が注目を集めた
できることが多様化	働き方の常識が徐々に変化 ・ 例：資料の目次作り、議論の観点出し、壁打ち相手として活用 ・ 例：GitHub Copilotを利用して効率的にコーディング GPT-4等のLLMに加え、生成AIをベースとしたクリエイティブ用途のモデルも増加 APIやプラグインによる機能拡張が一般的に
開発競争の加速・熾烈化	「独自にカスタマイズしたGPT」をノーコードで開発することが可能に (ChatGPTのGPTs) OpenAIのChatGPTは、高度な分析機能を備えたクラウド版「モデルアシスタント」をリリース AnthropicはClaude2を、GoogleはGeminiをリリース
AIガバナンスの重要性の再発見	国内ではMetaのLlama 2等をベースにした日本語強化モデルの模索が進捗 LLMの性能評価手法の検証や、新たな手法の提案が進捗 LLM等の生成AIの普及に伴い、実用段階ならではの問題が明らかに ・ 例：ハルシネーションや倫理的な問題、プロンプトインジェクション攻撃などの新たな脅威 ・ 例：なりすましの問題、プライバシーの保護、権利の問題や著作者の心情への配慮

これは1年前の話

たった1年と思えないほど

生成AIを取り巻く状況は様変わり

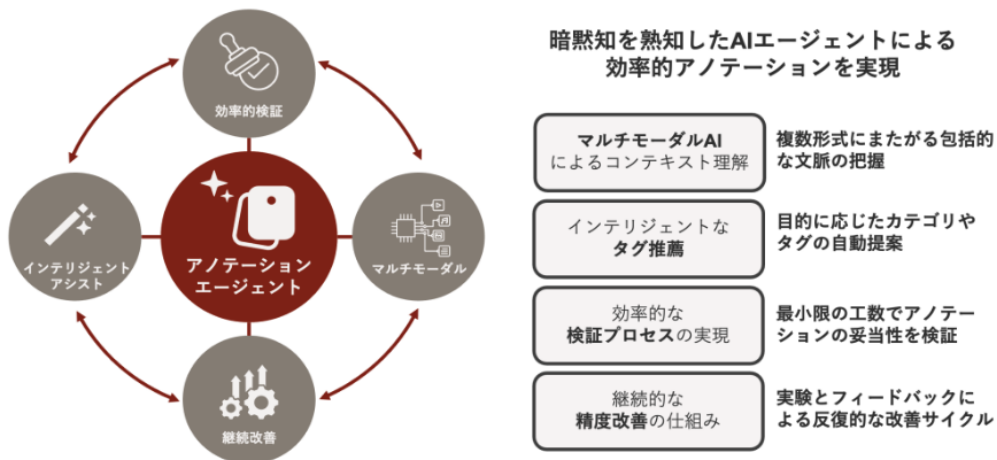
第2章：2025年の機械学習プロジェクト LLMをはじめとする生成AIの概況

生成AIの応用方法の一つの解である「自律的に働くAIエージェント」に注目が集まっています。

ブレインパッド、自律型AIエージェントサービス第一弾として、
アノテーションエージェントサービスを提供開始
－ マルチモーダルAIにより高度に文脈を理解、インテリジェントなタグ候補を提案 －

株式会社ブレインパッド(本社:東京都港区、代表取締役社長 CEO 関口 朋宏、以下:ブレインパッド)は12月12日、日本企業のデジタル競争力の向上および本質的なDXの実現を目指し、自律型AIエージェントサービスを順次発表していくこと、および、その第一弾としてアノテーション(*1)エージェントサービス「BrainPad アノテーションエージェント」を発表します。

当社は、2025年2月より「BrainPad アノテーションエージェント」の試験提供を開始します。



株式会社ブレインパッド 公式WEBサイト ニュースリリースより引用
<https://www.brainpad.co.jp/news/2024/12/12/22440>, (参照2025-01-07)

その他参考記事

岡本拓也. ブレインパッド、アノテーション特化のAIエージェント投入へ AI事業を数十億円規模に拡大狙う. EnterpriseZine. 2024.

<https://enterprisezine.jp/news/detail/21043>, (参照2025-01-07)

小河愛実・杜師康佑. 25年はAIエージェントが働く アポも資料作成も自律的に. 日本経済新聞. 2024.

<https://www.nikkei.com/article/DGXZQOUC13BJZ0T11C24A2000000/>, (参照2025-01-07)

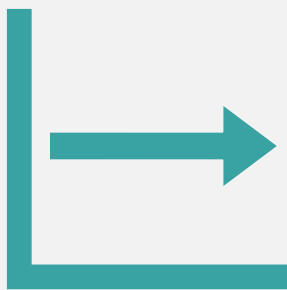
第2章：2025年の機械学習プロジェクト 変わるものと、変わらないこと

生成AI時代といっても、大切なことはさほど変わりません。



変わること

- 適用する技術（より進化し、できることが増加）
- 利用する道具（より抽象化されて便利に）
- 作業効率（生成AIの活用で効率向上、一部作業は代替可能に）
- 創造性を伴わないタスクの遂行能力の相対的な価値（同じことをしても価値は目減り）
- etc.



変わらないこと

- 自ら問いを立てて思考することや、しっかりとした言語化の重要性
- 意思決定や責任を負う機能の多くを人間が担うこと
- 人々の習慣・価値観を変化させる難しさ（工夫とパワーが変わらず必要）
- データサイエンスを取り巻く状況が「変化し続ける」こと
- etc.

質疑応答

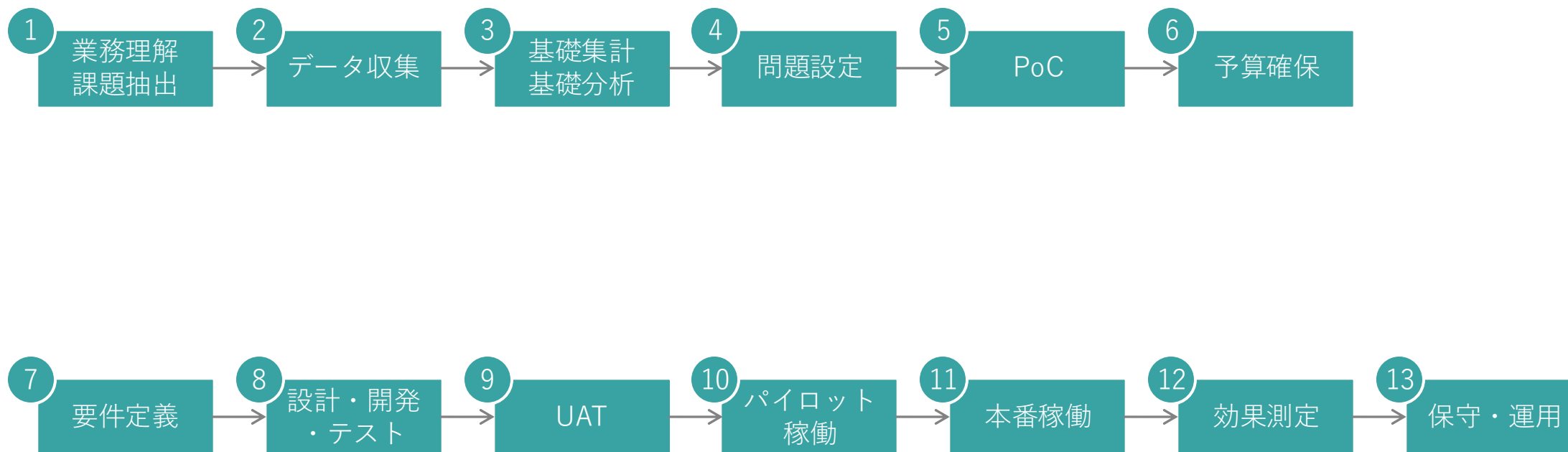
第3章

社会実装を阻む「罨」と その解決策

第3章：社会実装を阻む「罫」と、その解決策

一般的な機械学習プロジェクトのプロセス

第1章で触れた工程を細分化し、並べ直してみます。Kaggle^{*1}のように初めから「綺麗な問題」が用意されているわけではなく、必要なタスクが多岐にわたることがわかります。

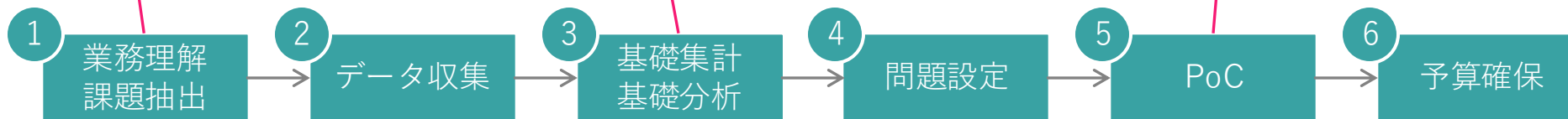


*1: 機械学習などデータに関連したコンペティションを提供するプラットフォーム。企業などの組織が課題とデータを投稿し、世界中のデータサイエンティストがより良い結果を目指してモデル構築に挑む
Kaggle公式Webサイト: <https://www.kaggle.com/>

第3章：社会実装を阻む「罨」と、その解決策 現実

前スライドで機械学習プロジェクトの膨大なタスクについて示しました。
しかし、それだけではありません。これらのプロセスには、たくさんの「罨」が待ち構えています。

その課題は本質か？
(課題感の偏り)



ふりだしに戻る
(報われない集計・分析)

進むも地獄、退くも地獄
(曖昧な出口戦略)

現実のデータは汚い！

(データが「印刷物」で「間違っている」ことも……)

問題設定は難しい

(できること ≠ 利益)

モデルの性能、どう測る？

(モデルの性能 ≠ 説得力)

様々な制約
(インフラ、政治等)



信頼を得るのは難しい
(利害の不一致)

学習データにない未来
(自然災害、どうする?)

順風満帆とは限らない
(継続の難しさ)

その開発、誰がやる？
(分析のプロ ≠ 開発のプロ)

思わぬところに考慮漏れ
(未知のデータのIF等)

ビジネスインパクト、どう測る？
(モデルの性能 ≠ 利益)

第3章：社会実装を阻む「罨」と、その解決策

【業務理解・課題抽出】その課題は本質か？

- 1 業務理解
課題抽出
- 2 データ収集
- 3 基礎集計
基礎分析
- 4 問題設定
- 5 PoC
- 6 予算確保
- 7 要件定義
- 8 設計・開発
・テスト
- 9 UAT
- 10 パイロット
稼働
- 11 本番稼働
- 12 効果測定
- 13 保守・運用

理想

効果の見込める施策が明らかであり、かつ組織内の大半が同じ方向を向いている

罨

- 上層部が掲げるDX構想と現場の実態に乖離があり、**地に足のついた取組みに繋がらない**
- 本質的（潜在的）な課題を認識できず、**対症療法的なアプローチに終始してしまう**
- コスト削減などの**守備的な課題へのアプローチが優先されてしまい、業務の在り方や行動様式の変化に繋がらない**
- **課題の言語化が不十分**で、解決の対象が曖昧

解決策

- ✓ 複数のレイヤー・役割の関係者にヒアリングした上で、現実的なことから取り組んでいく計画を立てる
- ✓ 課題だと聞いた内容を真に受けず、ヒアリング結果を踏まえて深く考察し、本質的な課題にたどり着く
- ✓ 目的ドリブンで取り組むべき課題について検討し、必要に応じて新たな付加価値を生むための「攻めの投資」をする
- ✓ 解決の対象となる課題を丁寧に構造化・言語化して解像度を上げ、関係者間で合意をとる

一言コメント

この段階で何を課題とするか？がプロジェクトのその後を左右します。
ビジョンや果たすべき社会的使命との整合性を保った上で、地に足のついた取組みを行うことが重要です。

第3章：社会実装を阻む「罫」と、その解決策

【データ収集】現実のデータは汚い！

- 1 業務理解
課題抽出
- 2 データ収集
- 3 基礎集計
基礎分析
- 4 問題設定
- 5 PoC
- 6 予算確保
- 7 要件定義
- 8 設計・開発
・テスト
- 9 UAT
- 10 パイロット
稼働
- 11 本番稼働
- 12 効果測定
- 13 保守・運用

理想

構造化されたデータがクラウド上の列指向DBに格納されており、容易に収集可能

罫

- 「利用できるデータがたくさんある」と聞いていたが、**そのデータは実は印刷物**であった
- データはあるが、**中身の数値や文字列が何を意味するかがわからない**
- データはあるが、データ保有部門との関係が悪く、**データを渡してもらえない**
- データはあるが、**常に上書き処理されており、蓄積されていない**

解決策

- ✓ 現状のデータの品質について、関係者一同で事前に認識を合わせる
- ✓ 必要なメタ情報を可視化した上で、現実との差分を把握し、情報の収集や提供依頼をする
- ✓ 役職者からトップダウンで業務命令が下るように関係者との調整に奔走する
- ✓ データを定期的に蓄積するスキームを作り、すぐにデータの蓄積を開始する

一言コメント

高度なデータ活用が想定されたデータはまだ少なく、データ整備を推進していく気概が求められます。また、昨今個人情報の取扱いが厳しくなりつつあることにも注意が必要です。（GDPRや改正個人情報保護法）

第3章：社会実装を阻む「罫」と、その解決策

【基礎集計・基礎分析】ふりだしに戻る

- 1 業務理解
課題抽出
- 2 データ収集
- 3 基礎集計
基礎分析
- 4 問題設定
- 5 PoC
- 6 予算確保
- 7 要件定義
- 8 設計・開発
・テスト
- 9 UAT
- 10 パイロット
稼働
- 11 本番稼働
- 12 効果測定
- 13 保守・運用

理想

データ分析基盤とデータ定義書が整備されており、偏りのない正確なデータが容易に利用可能

罫

- 一部のデータのみを受領しており、**母集団に偏りがある**
- ある程度分析を進めた後に、オペレーションミスと思しき**異常値の存在**が確認された
- **同一IDの商品が複数存在**するように見えるが、**データ定義書がなく理由がわからない**
- データの性質上、**合理的に補完（補間）することが不可能な欠損値の存在**が確認された

解決策

- ✓ 偏りのないデータを母集団として集計・分析できるように事前に手を打っておく
- ✓ 「初手・データを見る」を重要視する（データ収集の段階で実施しておくのが望ましい）
- ✓ テーブル間（あるいはファイル間）の関係性について推測しつつ、関係者に事実関係を確認する
- ✓ 欠損値の取扱い方針について関係者と協議し決定するとともに、影響について整理し、関係者間で共通認識にしておく

一言コメント

前提が誤っていれば、いかなる集計・分析結果も台無しです。
正しい解釈と作業の効率化のため、事前の準備に万全を期しましょう。

第3章：社会実装を阻む「罫」と、その解決策

【問題設定】問題設定は難しい

- 1 業務理解
課題抽出
- 2 データ収集
- 3 基礎集計
基礎分析
- 4 問題設定
- 5 PoC
- 6 予算確保
- 7 要件定義
- 8 設計・開発
・テスト
- 9 UAT
- 10 パイロット
稼働
- 11 本番稼働
- 12 効果測定
- 13 保守・運用

理想

解くべき問いと解ける問いが一致しており、特段疑問を抱かずに機械学習の問題に落とし込める

罫

- 賞味期限の長い「かんぴょう」の需要量を完璧に予測できるモデルが完成したが、**ビジネス的に意味があるとは到底思えない**
- **実はMAE等の指標よりも「大外し」をなくすことが重要**であることを後から指摘された
- 問題自体は解けるが、**解けば解くほど赤字が拡大**する見込みである
- 99%の精度を達成したが、**そもそも厳密に解ける必要があり**、実用化が見込めない

解決策

- ✓ 本当に解くべき問いは何か？について、必要なステークホルダーを巻き込んで議論する
- ✓ 機械学習の問題としての解きやすさと、ビジネス面の効果のバランスのとれた問題を設定する
- ✓ 現実的なコスト感で対応できる見通しが立つかを考える（スケーラビリティ等も重要な要素）
- ✓ 機械学習にこだわらず、解くべき問いに適した手法を選択する（ルールベース/数理最適化）

一言コメント

「解くべき問い」と「解ける問い」は往々にして一致しません。
時には機械学習以外の方法を検討する必要があります。（数理最適化等も課題解決のための強力な手段）

第3章：社会実装を阻む「罨」と、その解決策

【PoC】進むも地獄、退くも地獄

- 1 業務理解
課題抽出
- 2 データ収集
- 3 基礎集計
基礎分析
- 4 問題設定
- 5 PoC
- 6 予算確保
- 7 要件定義
- 8 設計・開発
・テスト
- 9 UAT
- 10 パイロット
稼働
- 11 本番稼働
- 12 効果測定
- 13 保守・運用

理想

特に躓くことなく予定通りにPoCが進捗し、実用化できることが誰の目にも明らか

罨

- 結果の良し悪しを判断できず、**出口が見えないまま**ずるずるとPoCを継続してしまう
- なかなか成果が出ず、いわゆる「**PoC死**」に至ってしまう
- **モデル自体の改善にのみ全力**を注いでしまう
- リッチな検証の仕組みを整えてPoCに挑んだが、**PoC終盤で大きな仕様変更**を余儀なくされる課題が見つかってしまい、対応できなくなった

解決策

- ✓ 事前にPoCのゴールを定量・定性の両面で定義し、期間を区切って評価する（比較対象を決めて相対評価するのも手）
- ✓ できないことがわかるのも一つの成果と捉え、解く問題やアプローチを再考する
- ✓ モデルそのものだけでなく、モデル周辺のあらゆる部分を工夫する（劇的に改善することがある）
- ✓ PoC期間中に新たな課題が見つかることを見越し、小回りの利く仕組みで複数回の検証を回す（Quick & Dirty）

一言コメント

ゴールを明確に定めてPoCに取り組み、結果が出なければ戦略的撤退を図るのも勇気です。

第3章：社会実装を阻む「罨」と、その解決策 【PoC】モデルの性能、どう測る？

- 1 業務理解
課題抽出
- 2 データ収集
- 3 基礎集計
基礎分析
- 4 問題設定
- 5 PoC
- 6 予算確保
- 7 要件定義
- 8 設計・開発
・テスト
- 9 UAT
- 10 パイロット
稼働
- 11 本番稼働
- 12 効果測定
- 13 保守・運用

理想

どこからどう見てもモデルの性能、およびビジネス上の効果に疑いの余地がない

罨

- 機械学習に過度の期待をされており、意思決定者の**期待を意図せず裏切ってしまう**
- 意思決定者が**性能指標を理解できない**、あるいはモデルの性能がビジネス上の利益に結びつく実感が湧かず**投資判断ができない**
- **予測値をどれくらい信用できるかがわからない**
- モデルの解釈性が低く、その**モデルを信頼する根拠として不足がある**

解決策

- ✓ 「良いモデル」を緻密に定義し、事前に役職者を含めて合意を取っておく
- ✓ 誰にでもわかりやすく、かつ本質を損なわない指標を定義した上でバックテストする
- ✓ 予測の不確かさを扱える手法を選択する
- ✓ 無理に深層学習に寄せず、回帰木などのモデルも候補に入れた上で、性能と解釈性のバランスをとる
- ✓ SHAP等で特徴量の寄与度の分析を試みる

一言コメント

基本的に、「交差検証しました。はいOK!」とはならないと考えておくべきです。
昨今は「説明可能なAI (XAI)」が注目を集めるなど、モデルの説明可能性に関心が高まっています。

第3章：社会実装を阻む「罨」と、その解決策

【要件定義】様々な制約

- 1 業務理解
課題抽出
- 2 データ収集
- 3 基礎集計
基礎分析
- 4 問題設定
- 5 PoC
- 6 予算確保
- 7 要件定義
- 8 設計・開発
・テスト
- 9 UAT
- 10 パイロット
稼働
- 11 本番稼働
- 12 効果測定
- 13 保守・運用

理想

ベストプラクティスに基づき、一般的に良しとされる要件を過不足なく定義していける

罨

- 既存のデータ基盤との兼ね合いやセキュリティ上の制約により、**クラウドコンピューティングへのスムーズな移行が難しい**
- **既存の仕組みを変えたくない部署**の発言権が社内で強い
- **要件がオーバースペックで赤字**を見込んでしまう
- **個別要件の対応**に追われ、**横展開が難しくなる**

解決策

- ✓ 既存の基盤をある程度活かした仕組みを構築し、段階的に移行していく計画を立てる
- ✓ 役職者を巻き込み、定期的にディスカッションの場を設けて意思決定を促す
- ✓ 最終的に達成したいことから逆算し、要件を取捨選択。見送った要件は後続フェーズで検討する
- ✓ 横展開を見越して、汎用的な要件と個別要件を分けて管理・開発する

一言コメント

本番稼働を目指すとなると、個別の事情と汎用性のバランスを意識して要件を定義していく必要があります。「理論の理解」や「実装力」では太刀打ちできない領域もあります。得意な人に任せてしまうのも手です。

第3章：社会実装を阻む「罫」と、その解決策 【設計・開発・テスト】その開発、誰がやる？

- 1 業務理解
課題抽出
- 2 データ収集
- 3 基礎集計
基礎分析
- 4 問題設定
- 5 PoC
- 6 予算確保
- 7 要件定義
- 8 設計・開発
・テスト
- 9 UAT
- 10 パイロット
稼働
- 11 本番稼働
- 12 効果測定
- 13 保守・運用

理想

データサイエンスとエンジニアリングの両方に長けた人材が、プロジェクトを一貫して主導する

罫

- PoCが終わり、いよいよ開発フェーズとなったものの、いざ本番稼働させるシステムを開発するとどのように開発していけばいいかわからない
- データサイエンスに長けたメンバーとエンジニアリングに長けたメンバーがそれぞれいるが、**コミュニケーションに難があり両車輪が動かない**
- PoCで書いたコードを本番環境に移植したいが、**中途半端に抽象化されており取扱いに困る**

解決策

- ✓ データサイエンスの担当者の他に、機械学習まわりのエンジニアリングの担当者（機械学習エンジニア等）をアサインする（可能であれば初期段階からアサインしておき、スムーズに本番稼働させるシステムの開発に入れるように準備しておく）
- ✓ チーム内で最低一人が「翻訳者」となり、メンバー間のコミュニケーション促進の役割を担う
- ✓ PoCの段階からシステムリリースを見越してクラス設計等を丁寧にしておくか、あるいは敢えてJupyter Notebookで書き下す以上のことをしない

一言コメント

データサイエンス分野の成熟に伴い専門分化が進んでいる側面と、実務上の必要性からエンジニアリングに長けた人材が増えている側面の両方があります。（現場所感）状況に応じてベストな編成・役割分担を。

第3章：社会実装を阻む「罨」と、その解決策

【UAT】信頼を得るのは難しい

- 1 業務理解
課題抽出
- 2 データ収集
- 3 基礎集計
基礎分析
- 4 問題設定
- 5 PoC
- 6 予算確保
- 7 要件定義
- 8 設計・開発
・テスト
- 9 UAT
- 10 パイロット
稼働
- 11 本番稼働
- 12 効果測定
- 13 保守・運用

理想

モデルの性能が良く、現場からの評判も上々。スムーズに次のフェーズに移行できる

罨

- 性能の良いモデルを提供しても**現場担当者には旨味がなく**、既存のオペレーションを変えたくない層から**ネガティブな意見が出る**
- 予測が外れたごく一部について現場担当者に固執されてしまい、**モデルを信頼してもらえない**
- **やらされている感**や、利用者の**プライドを傷つける**ことに繋がってしまう
- 確かにモデルの性能は良いが、実際に現場のオペレーションに組み込んでみたところ、**使いにくい部分があることがわかった**

解決策

- ✓ 予測が当たった場合のメリットについて、経営目線だけでなく、現場目線で整理する
- ✓ 予測が外れた原因を可能な範囲で分析し、説明して納得してもらう
- ✓ 「選択の自由」を残し、最終的な意思決定を利用者に委ねるサービス仕様を検討する
- ✓ UI/UXの設計・開発工数を確保し、システム利用時のハードルを下げる
- ✓ ユーザーからの意見を漏れなく吸い上げ、改善すべき点については改善を試みる

一言コメント

利用者に「使う側のメリット」を提示し、Win-Winの関係でプロジェクトを進めていくのが正解です。いかに高度なモデルも、結局のところ使ってもらえなければ宝の持ち腐れです。

第3章：社会実装を阻む「罨」と、その解決策

【パイロット稼働】思わぬところに考慮漏れ

- 1 業務理解
課題抽出
- 2 データ収集
- 3 基礎集計
基礎分析
- 4 問題設定
- 5 PoC
- 6 予算確保
- 7 要件定義
- 8 設計・開発
・テスト
- 9 UAT
- 10 **パイロット
稼働**
- 11 本番稼働
- 12 効果測定
- 13 保守・運用

理想

PoCと同様の性能が出ており、目立ったバグもなく順調に稼働している

罨

- **PoC時と同等の性能が示せず、本番稼働に踏み切れない**
- マスターの内容が変化しており、**存在しないはずのカテゴリカル変数が特徴量として投入されてしまう**
- 定められた時刻までに必要なデータがIFされてこず、**必要な特徴量がNULLのまま予測処理が行われてしまう**
- 連携された**データの不備が原因で不具合が生じたにもかかわらず、モデルの責任にされてしまう**

解決策

- ✓ PoC時と同じ品質のデータが使えるとは限らないことを認識し、可能な限り手を打っておく
- ✓ 本番に近い形でバックテストを実施し、ある程度の性能が出ることを担保しておく
- ✓ 機械学習モデルによる予測値を過信せず、セーフティネットとして異常値を回避するための仕組みを複数用意しておく
- ✓ 各関係者が責任を持つべき領域（責任分界点）をあらかじめ明確にしておく

一言コメント

たった一度の失敗で、機械学習モデルのような「わかりにくいもの」に対する信頼は崩れ去ります。そうならないために、「事故が起きない仕組みづくり」を徹底する必要があります。

第3章：社会実装を阻む「罨」と、その解決策

【本番稼働】学習データにない未来

- 1 業務理解
課題抽出
- 2 データ収集
- 3 基礎集計
基礎分析
- 4 問題設定
- 5 PoC
- 6 予算確保
- 7 要件定義
- 8 設計・開発
・テスト
- 9 UAT
- 10 パイロット
稼働
- 11 本番稼働
- 12 効果測定
- 13 保守・運用

理想

システムは順調に稼働。社会情勢に大きな混乱はなく、モデルは質の高い予測をし続けている

罨

- **学習データの期間に存在しないイベント**が行われることとなった
- **観測史上最大の台風**が日本列島に上陸し、猛威を振るう見込みである
- **突然のパンデミック**。モデルはパンデミック時の需要量の傾向を学習しておらず、妥当な水準の予測ができる保証がない

解決策

- ✓ 解釈性の高いモデルやルールベースのアルゴリズムとの2段構えの仕組みを予め構築し、必要に応じてスイッチできるようにしておく
- ✓ 緊急時に運用回避できるように、緊急時用のオペレーションを組み、日頃から周知しておく
- ✓ 現場の状況を丁寧にヒアリングしつつ、モデルの利用可否や再開タイミングについて一つひとつ判断する

一言コメント

コロナ禍に限らず、経験したことのない出来事が起こる可能性は多くの場合に存在します。想定外の状況に備え、システムや運用の設計に反映しておくことが求められます。

第3章：社会実装を阻む「罫」と、その解決策

【効果測定】ビジネスインパクト、どう測る？

- 1 業務理解
課題抽出
- 2 データ収集
- 3 基礎集計
基礎分析
- 4 問題設定
- 5 PoC
- 6 予算確保
- 7 要件定義
- 8 設計・開発
・テスト
- 9 UAT
- 10 パイロット
稼働
- 11 本番稼働
- 12 効果測定
- 13 保守・運用

理想

モデルは期待を上回る性能を示し続けており、具体的なアクションにより利益に繋がっている

罫

- モデルの性能は高い値を示しているが、それが具体的にどのように**ビジネスに貢献できているかがわからない**
- モデルの性能とビジネスへの貢献度に**相関はあるが、因果がわからず効果を測れない**
- モデルによる予測結果が**具体的なアクションに繋がっていない**

解決策

- ✓ モデルの性能が必ずしもビジネス上の効果に結びつかないことを認識し、ビジネス上のインパクトを定量的に可視化する指標を新たに作成する
- ✓ A/Bテストや統計的因果推論などにより、アクションとビジネスへの貢献度の因果関係を明らかにする
- ✓ モデルによる予測結果を意思決定などのアクションに確実に繋げる「出口」を用意する（数理最適化なども選択肢の一つ）

一言コメント

ビジネス上の効果が示せなければ、次の投資判断にダイレクトに響いてきます。そうならないために手を打つことが、機械学習の「社会実装」の拡大に繋がります。

第3章：社会実装を阻む「罨」と、その解決策

【保守・運用】順風満帆とは限らない

- 1 業務理解
課題抽出
- 2 データ収集
- 3 基礎集計
基礎分析
- 4 問題設定
- 5 PoC
- 6 予算確保
- 7 要件定義
- 8 設計・開発
・テスト
- 9 UAT
- 10 パイロット
稼働
- 11 本番稼働
- 12 効果測定
- 13 保守・運用

理想

ローンチ後、特に問題なくモデルを運用できており、システム改修の必要性も特段ない

罨

- データ取得元のテーブル定義に認識していない変更があり、**予測前処理時にエラーを吐いてしまう**
- 特徴量として使っていた**データが諸般の事情で使えなくなる**（データ提供元の方針変更、組織の意向等）
- 日々利用している**モデルの性能が徐々に下がってしまう**
- 機械学習の運用に強いチームが存在せず、**長期的な運用に適したチーム・人材が確保できない**

解決策

- ✓ 変更情報をキャッチできないことがないように、ステークホルダーとの情報共有のスキームを作っておく
- ✓ 複数のリスクの高いデータの使用を避ける
- ✓ モデルの再学習や、新たな学習済モデルへの差し替えの仕組みをシステムや運用に組み込んでおく
- ✓ 複雑かつ解釈性の低いモデルの利用を避け、運用の難易度を下げる
- ✓ 運用タスクのテンプレ化・抽象化を進め、高度な専門性を必要としない領域を拡大する

一言コメント

機械学習プロジェクトでも、“負の遺産”を残さないようにシステム・運用を設計することが必要不可欠です。また、時間の経過と共に変わりゆくトレンドを捉え続けるための仕組みづくりも必要です。

第4章

AI大活用時代に 何ができるか

第4章：AI大活用時代に何ができるか データサイエンティストに求められる素養

AI大活用時代にあっても、社会のアップデートを目指す私たちに求められる素養*1は変わりません。

基本の道具を使いこなせるか	理系学部レベル以上の数学や統計学の知識と応用力、あるいはこれらの獲得のためのポテンシャルと意欲
外部環境の変化に 追従できるか	変わりゆく外部環境を受け入れて自己研鑽を継続していく心構え
	利用経験のない技術やツールを敬遠せずに試す習慣
ビジネスの場での 価値創出を目指せるか	ビジネスの場における価値創出への直接的な貢献がミッションであると理解し、それを目指していけること
	どのような相手でも臆せずにコミュニケーションが取れること
自身の力で 考え抜くことができるか	思考力と応用力を備えていること
	前例にとらわれず、自ら考え解決策を導き出す力
	必要な対策を事前に施すためにあらゆる状況を想像する力
データをリスペクトし 情熱を注げるか	想定外の困難を打破する胆力とラストマンシップ
	データと向き合う時間を苦にせず、情熱を持ち続けられること
	データが語る事実を曲げずに結果を受け入れ、誠実に対応する態度

*1: 詳しくはこちらのブログ記事を。内池もえ・兵藤誠・川崎悠介. 【社員が解説】データサイエンティストとは？仕事内容やAI・DX時代に必要なスキル. DOORS DX Media BY BrainPad. 2023. https://www.brainpad.co.jp/doors/knowledge/01_about_data_scientist/, (参照2025-01-07)

第4章：AI大活用時代に何ができるか 特に気に留めてほしいこと

スキルを磨くのはもちろんのこと、マインド面も大切にしつつ、新たな時代をつくっていただきたいです。

目的 思考

- ✓ 目的を強く意識し、目的に沿った判断・行動をする（本来の目的から逸れる引力に屈しない）

メタ 認知

- ✓ 自らの強みと弱みを客観的に捉え、強みを活かせる可能性を模索する

学び 続ける

- ✓ 変化の速い時代に取り残されないように、いつからでも、いつまでも学び続ける

時代を つくる

- ✓ 前例にとらわれず、新たな価値の創造に挑む（∵機械学習プロジェクトは世界初の挑戦ばかり）

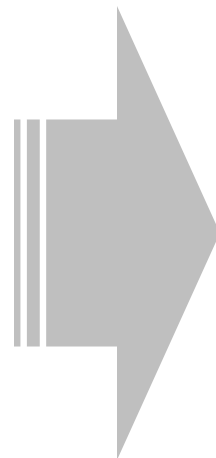
第4章：AI大活用時代に何ができるか 持てる力をフル活用せよ！

これまでの学び、経験、感じたことを総動員して「新たな価値の創出」に挑戦してみませんか。

これまで

- 3つの力を駆使してデータを活用
- 3つの力は相互に補完し合う関係
- 3つの力で許される場面が多かった

ビジネス力
データサイエンス力
データエンジニアリング力



これから

- 3つの力はそれぞれ更に磨いていく
- 3つの力は**相互補完的ではなく、掛け合わせる**ことが前提
- 3つの力に留まらず、その補集合の**ありとあらゆるものが次代を拓く力**に

Complement

3つの力に統計学や機械学習以外の専門性を掛け合わせたり、その人ならではのバックグラウンドや強みを活かしたりする

重なる領域を広げる

ビジネス力
データサイエンス力
データエンジニアリング力

参考：一般社団法人データサイエンティスト協会. 2019年度スキル定義委員会活動報告. 2019. p.6.
https://www.datascientist.or.jp/symp/2019/pdf/1115-1155_skill.pdf, (参照2025-01-07)

まとめ

本日本日お伝えしたかったこと

まとめ
本日お伝えしたかったこと



なぜやるのか

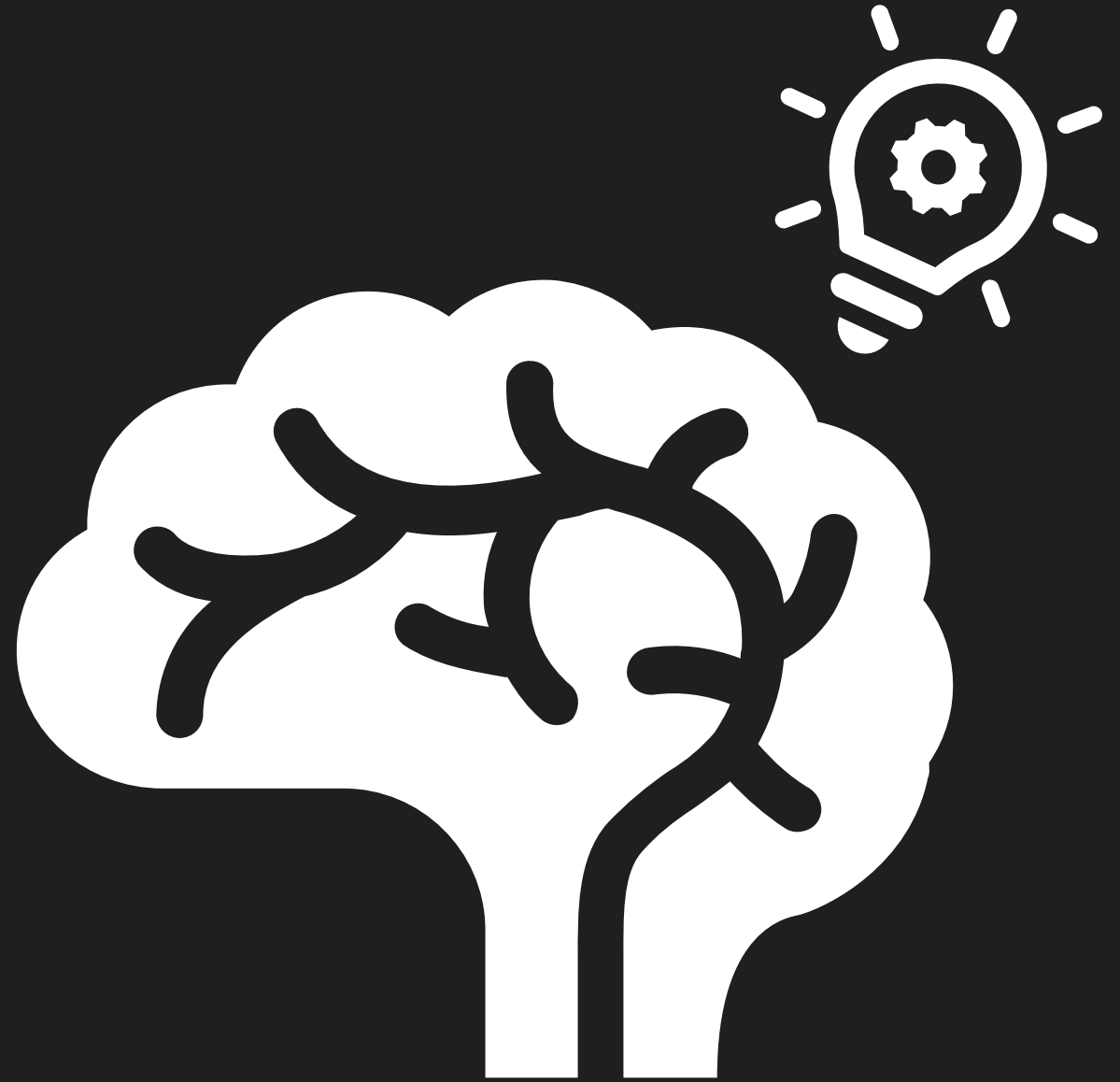


何に立ち向かうのか



どうやるのか

思考せよ
創造せよ



関連情報

内池もえ・兵藤誠・川崎悠介. 【社員が解説】データサイエンティストとは？仕事内容やAI・DX時代に必要なスキル. DOORS DX Media BY BrainPad. 2023.

https://www.brainpad.co.jp/doors/knowledge/01_about_data_scientist/, (参照2025-01-07)

DOORS編集部. AIを社会実装するためにデータサイエンティストができること. DOORS DX Media BY BrainPad. 2024.

https://www.brainpad.co.jp/doors/contents/social_implementation_of_ai_data_scientist/, (参照2025-01-07)

国立研究開発法人産業技術総合研究所. 機械学習品質マネジメントガイドライン 第3版. 2021.

<https://www.digiarc.aist.go.jp/publication/aigm/guideline-rev3.html>, (参照2025-01-07)

経済産業省. 「生成AI時代のDX推進に必要な人材・スキルの考え方」を取りまとめました. 2023.

<https://www.meti.go.jp/press/2023/08/20230807001/20230807001.html>, (参照2025-01-07)

経済産業省. AIガバナンス. 2021.

https://www.meti.go.jp/policy/it_policy/ai-governance/index.html, (参照2025-01-07)

経済産業省. AI事業者ガイドライン（第1.0版）. 2024.

https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/20240419_report.html, (参照2025-01-07)

個人情報保護委員会. 改正個人情報保護法 特集. 2020.

https://www.ppc.go.jp/news/kaiseihou_feature/, (参照2025-01-07)

大城信晃・マスクド・アナライズ・伊藤徹郎・小西哲平・西原成輝・油井志郎. AI・データ分析プロジェクトのすべて. 技術評論社. 2020.

<https://gihyo.jp/book/2021/978-4-297-11758-0>

ゆずたそ・渡部徹太郎・伊藤徹郎. 実践的データ基盤への処方箋——ビジネス価値創出のためのデータ・システム・ヒトのノウハウ. 技術評論社. 2021.

<https://gihyo.jp/book/2021/978-4-297-12445-8>

森下光之助. 機械学習を解釈する技術～予測力と説明力を両立する実践テクニック. 技術評論社. 2021.

<https://gihyo.jp/book/2021/978-4-297-12226-3>

安井翔太. 効果検証入門——正しい比較のための因果推論／計量経済学の基礎. 技術評論社. 2020.

<https://gihyo.jp/book/2020/978-4-297-11117-5>

岡野原大輔. 大規模言語モデルは新たな知能か——ChatGPTが変えた世界. 岩波書店. 2023.

<https://www.iwanami.co.jp/book/b625941.html>

安野たかひろ（東京都知事候補）公式ホームページ. 2024.

<https://takahiroanno.com/>, (参照2025-01-07)

(株)ブレインパッド 有志. OpenBrainPad Project.

社内にある技術資料の公開等を行っています

<https://brainpad.github.io/OpenBrainPad/>, (参照2025-01-07)

Thank you for your time and attention!

Appendix

過去の課題

※余力があれば是非考えてみてください

アイスブレイク DXを構想せよと言われたら (1/2) – 課題

以下の状況を踏まえ、皆さんならどのような構想を練るかを考えてアウトプットしてみてください。

- 東証プライム市場に上場する大手製造業A社は、残念ながらDXの潮流に乗り遅れています。
- 競合他社は既に積極的にDXに取り組んでいますが、A社では全社横断DX組織が立ち上がったばかりです。
- そんな中、A社経営陣はDXを強力に推し進める方針を固めました。
- あなたはA社のIT戦略の最高責任者として、ただちにDXの構想を練る必要があります。
- DXの予算は潤沢です。しかし、必ずやDXを達成し成果を出さなければなりません。

皆さんならどのようなことに、どのような順番で取り組む構想を練りますか？

3分間で考えてみてください

この問題に明確な答えがあるわけではありません。ですが、あらゆる視点で考えなければなりません。

DXの目的の視点

- 何のためにDXを達成しようとするのか
- 組織内の強化 or 顧客体験価値向上 or 次世代標準を示し業界/社会をアップデート のどれを目指すか
- DXを通じて自社のポジションや社会的役割をどう位置づけるか

ビジョンと戦略の視点

- 3年後、5年後、10年後はどのような状態に達しているか、そのときの社会の状態はどうか
- 何から始め、どのような道筋を辿って発展させていくか (DX展開ロードマップ)
- 他社との差別化をどう図るか

方法論の視点

- どのような技術/ツールで何をどう解くか

投資対効果やガバナンスの視点

- 投資に見合った価値を創出できるか、その価値をどう測るか
- セキュリティ等のリスクとどう向き合うか

組織の視点

- 社員を育成して内製で取り組むか、外部の力を借りるのか
- DXの中核を担う人材をいかに確保し、長く活躍してもらうか (人事戦略と制度改革)
- 社内の文化と社員の価値観をいかにシフトさせていくか

生成AIの業務利用推進チームに配属されたら、どうする？ (1/2)

以下の状況を踏まえ、皆さんなら何を考え、どのようなことを実践するかを理由と共にアウトプットしてみてください。

状況

- 製造業を営むA社では、古くから伝統的な統計手法を用いたQC (品質管理) に力を入れていました。
- 一方で、昨今の機械学習等の新たな手法を取り入れることにはあまり積極的ではありませんでした。
- しかし、時代が進み状況が一変します。生成AIの世界的な盛り上がりや競合B社のDX成功事例を受けて、A社は自社の業務に積極的に生成AIを活用していく方針を固めました。
- そんな折、あなたはデータサイエンティストとしてA社に中途採用され、生成AIの業務利用推進チームに配属されました。
- 周囲に経験豊富なデータサイエンティストはおらず、どうやら実質的なリーダーとして周囲をリードしていく必要があります。

皆さんなら何を考え、どのようなことを実践しますか？

3分間で考えてみてください

生成AIの業務利用推進チームに配属されたら、どうする？ (2/2)

この問題に明確な答えがあるわけではありません。ですが、ありとあらゆることを考えなければならないことがわかります。

考えるべきこと (例)

- **何をどう解決するのか**
 - そもそも、解決したい業務上の課題は何か
 - 製造業における生成AI適用の先行事例/類似事例はあるか
 - 作業効率の底上げ等の守りに徹するか、それとも攻める姿勢をとるか
 - 生成AIの業務利用は目的達成のためのベストな手段か
- **実現可能か**
 - 現時点における生成AIの性能で実現可能か
 - 学習やチューニングに利用するデータを確保できるか
 - 社内のケイパビリティは十分か (技術力、ドメイン知識、実行力)
 - 生成AIの活用を検討・推進する人材を確保 (育成、採用 等) できるか
- **効果は出るか**
 - 投資に見合ったリターンが得られるか (金銭面、その他)
 - 効果をどのように測定するか、できるか
 - 利用者側のリテラシーは十分か、いかに啓蒙活動をしていくか
 - 生成AIの利用によって発生しうるリスクは想定/コントロール可能か
- **何をすべきか**
 - 関係者に対するヒアリングによる課題感の把握
 - 解決したい課題のしっかりとした構造化・言語化
 - 目指す姿や実現可能性に関する対話と期待値調整
 - 部署を跨いだプロジェクトチームの立ち上げ

我々が立ち向かわなければならないのは、まさにこのような問題の数々！

考えてみよう：パンデミック時の需要予測、どうする？ (1/2)

まずは1つ課題を出します。以下の状況を踏まえ、考えたことをアウトプットしてみてください。

状況

- 皆さんは、世界中に店舗を展開している和食チェーンにおいて、機械学習による需要予測プロジェクトのPMを務めています。
- このプロジェクトでは、翌月に必要になる食材の需要量をモデルで予測し、その予測結果をもとに食材が発注・納品されることを目指します。
- 既に皆さんは数々の試練を乗り越え、いよいよ本番稼働というタイミングになりました。
- ところが、このタイミングでパンデミックが起きてしまいました。このパンデミックはいつ收拾がつくかわかりません。
- 学習データは過去3年分しかなく、かつ過去3年間に類似のパンデミックは起こっていません。

皆さんならPMとして、この問題にどう立ち向かいますか？

3分間で考えてみてください

考えてみよう：パンデミック時の需要予測、どうする？ (2/2)

いかがでしたでしょうか。この問題に明確な答えがあるわけではありません。ですが、ありとあらゆることを考えなければならないことがわかります。この事実に気づくことが機械学習の「社会実装」に直結します。

思いつくこと (例)

- そんな時期に予測が当たるはずがないじゃないか！
- そもそも店舗は開いているのか？
- 需要予測が「大外れ」した場合の経済的損失やフードロスは？
- その責任は一体誰が取るのか？
- そもそもこの期間にモデルを稼働させるのか？
- 仮に稼働させるとして、予測値は「後処理」するべきではないか？
- その「後処理」は何が適切か？ルールベースなのか？
- ローンチを遅らせてみるのはどうか？
- ローンチを遅らせたとして、パンデミックの期間のデータはモデルに学習させていいのか？

我々が立ち向かわなければならないのは、まさにこのような問題の数々！

考えてみよう：モデルの品質、どこまで保証できる？ (1/2)

以下の状況を踏まえ、考えたことをアウトプットしてみてください。

状況

- 皆さんは法人向けに受託分析サービスを提供する企業に勤務しており、機械学習による与信審査プロジェクトのPMを務めています。
- プロジェクトのミッションは、あらゆるデータを活用し、より良い与信審査システムを構築することです。
- クライアントの金融機関における、過去の顧客のID、年齢、性別、居住地、雇用形態、勤続年数、家族構成、貸し倒れの有無などに加え、有償の市況データなども収集し、これらのデータを利用して貸し倒れリスク予測モデルを構築しました。
- 幸いなことに、PoC時に構築したモデルの予測性能がそれなりの水準に達しているように見えたため、ステアリングコミティで検証結果を報告したところ、システムリリースを目指してプロジェクトを前進させるように依頼されました。
- プロジェクトはクライアント企業の社長直轄で、「最終的に精度90%を達成すること」「品質が保証されること」「SLA*の提示」を強く求められています。

皆さんなら、どのように判断し、どのように行動しますか？

3分間で考えてみてください

*: Service Level Agreementの略。提供するサービスの品質保証レベルに関する合意事項。

考えてみよう：モデルの品質、どこまで保証できる？ (2/2)

いかがでしたでしょうか。この問題に明確な答えがあるわけではありません。ですが、ありとあらゆることを考えなければならず、思考停止させてもらえないことがわかります。皆さんはどこまで想像できたでしょうか？

思いつくこと (例)

- 精度90%の根拠はあるのか？達成すると何が嬉しいのか？
- 精度とは何を指しているのか？*
- データさえ増やしていけば改善され続けると誤解されていないか？
- 機械学習モデルの精度を保証するのは現実的なのか？
- 検証時の前提条件は本番環境で満たせるのか？
- 有償の市況データを調達し続けられる保証はあるか？
- モデルの検証結果に嘘はないか？ (想定外のリーク)
- この先市況が大きく悪化した場合、モデルの予測性能に再現性はあるのか？
- 予測性能そのもの以外にも、可用性や処理速度、モデルの解釈性や公平性などが保証対象となり得るのではないか？
- 動作保証ならできる可能性があるが、そのためには前提条件や免責事項を明示する必要があるのではないか？
- 性別を特徴量とする与信審査システムは公平性を欠いており、社会的要請を満たしていないのではないか？

我々が立ち向かわなければならないのは、まさにこのような問題の数々！

*: 二値分類の正解率、適合率、属性別の貸し倒れ率の誤差、これらに貸し倒れ時の損害規模で重みづけする必要の有無、あるいは属性別の貸し倒れリスクをある程度の幅で予測できればいいのか 等が考えられる。

考えてみよう：このプロジェクトは推進するべきか？（1/2）

以下の状況を踏まえ、プロジェクトを推進するべきかと、その結論に至った理由（または評価観点）をアウトプットしてみてください。結論が出せない場合は、どのような情報が不足しているかを列挙してみてください。

状況

- 皆さんは、データ分析や機械学習を扱う受託分析企業*1に勤めています。
- ここ数年間、日本全国に店舗を展開する和食チェーンから案件を受託し、機械学習による需要予測プロジェクトのPMを務めてきました。
- このプロジェクトでは次のタスクに取り組み、モデルや機能のほとんどを仕上げることができました。
 1. 翌月に必要となる食材の需要量を予測するモデルの構築
 2. モデルによる予測結果をもとに食材が発注・納品されるシステムの開発
- ところが、ローンチを間近に控えたある日、突如訪れたパンデミックにより、プロジェクトを凍結させることになりました。
- その後、2年の月日が流れてパンデミックが収束に向かい始めたタイミングで、和食チェーンからプロジェクト再開の打診を受けました。

和食チェーンからの打診を受け入れ、再びプロジェクトを推進していくべきでしょうか？

3分間で考えてみてください

*1: 顧客企業に伴走し、データや機械学習などに関する顧客企業の課題を解決することを主な事業内容とする企業

考えてみよう：このプロジェクトは推進すべきか？ (2/2)

この問題に明確な答えがあるわけではありません。ですが、ありとあらゆることを考えなければならないことがわかります。

観点

- そもそも社会的意義はあるか、社会を変革しうるか
- 機会費用を支払う価値はあるか (もっと有益な取組みはないか)
- パンデミック前後のデータで上手く学習できるか
- プロジェクトを凍結している間により良い手段が出てきていないか
- 凍結当時のアーキテクチャや運用設計を流用できるか、すべきか
- 改めて開発・運用体制を組み、維持していけるか
- 改正個人情報保護法などの法規制の影響を受けないか
- 和食チェーンのインバウンド需要はどうか (これまでとこれから)
- 昨今の国際情勢 (ロシア・ウクライナ戦争 等) が影響しうるか
- 円安傾向やインフレ、低水準の実質金利がどう影響するか
- マネタイズ可能か (ビジネスモデルとして優れているか)
- 自社のケイパビリティやブランド力を高めうるか
- 参画するメンバーにとってプラスになりうるか
- 横展開しうるか (知財の取扱いは適切か、システムとして汎用的か)

我々が立ち向かわなければならないのは、まさにこのような問題の数々！

考えてみよう：どのような提言/提案をするべきか？ (1/2)

以下の状況を踏まえ、クライアントA社に対してどのような提言/提案をするべきかを考えて、理由と共にアウトプットしてみてください。結論が出せない場合は、どのような情報が不足しているかを列挙してみてください。

状況

- 皆さんは、データ分析や機械学習を扱う受託分析企業*1に勤めています。
- つい先日、大手製造業のクライアントA社から以下の要望を受けました。
 - データならたくさんあるので、何らかの機械学習プロジェクトを企画・推進し、A社社員と協力して成果を出してほしい
 - ゆくゆくはA社社員にスキルトランスファーし、同様のプロジェクトに内製で取り組めるようにしてほしい
- A社は特定分野の製品の製造において国内外から高く評価されていますが、いわゆるDXの波には乗り遅れており、同業他社と比べるとデータの利活用が進んでいない状況です。
- 皆さんの所属企業とA社の取引は今回が初めてで、不特定多数に公開されている以上の情報はまだ得られていません。

皆さんなら、A社に対してどのような提言/提案をしますか？

3分間で考えてみてください

*1: 顧客企業に伴走し、データや機械学習などに関する顧客企業の課題を解決することを主な事業内容とする企業

考えてみよう：どのような提言/提案をするべきか？ (2/2)

この問題に明確な答えがあるわけではありません。ですが、ありとあらゆることを考えなければならないことがわかります。

考えるべきこと (例)

- データならたくさんあるというが、品質はどうか
 - ただちに有効利用できる状態なのか
 - データ基盤は整備されているか
 - データのメタ情報は十分で、適切に管理されているか
 - データに関する不明点を解消する方法は明らかか
- A社が本当に達成したいことは何か
 - 機械学習PJや内製化のその先の目的は何か
 - 機械学習PJや内製化は目的達成のためのベストな手段か
 - 抽象的な要望をいかに具体化していくか
 - どこまで本気か、いくらかけられるのか
- 内製化は必要か、そもそも可能なのか
 - A社にとってのメリット/デメリットは何か
 - 内製化を可能にするための人材や仕組みは揃っているか
 - A社と自社でどのように役割分担していくか
 - スキルトランスファーは本当に可能なのか、終わりはあるのか
- 必要な情報を集めることはできないか
 - A社の現時点におけるケイパビリティはどうか
 - 構想はどこまで理解されているか、部門間に利害はないか
 - ヒアリングを実施できないか
 - 短期間の分析トライアルで実態を把握できないか

我々が立ち向かわなければならないのは、まさにこのような問題の数々！

考えてみよう：モデルの性能、どう測る？ (1/2)

以下の状況を踏まえ、自分なりの答えを出してみてください。

状況

- 皆さんは、世界中に店舗を展開している和食チェーンにおいて、機械学習による需要予測プロジェクトのPMを務めています。
- 今まさに、解くべき問題を「来月必要な食材の需要量」と設定し、PoCを回そうとしています。
- 食材は米、野菜、肉、かんぴょうなど様々です。
- A国では「かんぴょう巻」が絶大な人気を誇っていますが、B国ではあまり人気がありません。

皆さんなら、どのような指標・考え方でモデルを評価しますか？

3分間で考えてみてください

考えてみよう：モデルの性能、どう測る？ (2/2)

いかがでしたでしょうか。この問題に明確な答えがあるわけではありません。ですが、ありとあらゆることを考えなければならないことがわかります。この事実気づくことが機械学習の「社会実装」に直結します。

恐らく皆さんが考えたこと

- 予測値と実績値の誤差を最小化すればよいのだから、素直にMAEで評価すればいいのではないかな？
- いや、「大外れ」は賞味期限の問題で修正がきかないのだからRMSEで評価するべきなのではないかな？
- 「当てるべきもの」と「当てなくていいもの」が存在するのではないかな？ (例えば、かんぴょうの需要量を当てても意味がない)

私ならこういうことも考える

- 「過剰予測」と「過小予測」にどのように重みづけをするかな？ (過剰在庫と販売機会損失の重みを天秤にかける)
- 国別、あるいは地域別に必要とするモデルの振る舞いは異なるのではないかな？
- 最終的に「良いモデル」であることをどう定義し、どう証明するかな？ (絶対評価とするかな？何かと比べて相対評価とするかな？それぞれの場合の効果試算をどのように行うかな？)

我々が立ち向かわなければならないのは、まさにこのような問題の数々！

考えてみよう：オリンピックイヤーの需要予測、どうする？ (1/2)

以下の状況を踏まえ、自分なりの答えを出してみてください。

状況

- 皆さんは、世界中に店舗を展開している和食チェーンにおいて、機械学習による需要予測プロジェクトのPMを務めています。
- 様々な困難を乗り越え、ようやくシステムローンチすることができました。
- ひとまず問題なく動いており、現場からの評判も上々です。
- しかし、100店舗を展開しているヨーロッパのA国で、来年オリンピックが開催されることに気づきました。
- 学習データは3年分しかなく、オリンピック期間の需要量については見当がつきません。
(ここではオリンピックに準ずる規模のイベントもなかったと仮定します)

皆さんならPMとして、この問題にどう立ち向かいますか？

3分間で考えてみてください

考えてみよう：オリンピックイヤーの需要予測、どうする？ (2/2)

いかがでしたでしょうか。この問題に明確な答えがあるわけではありません。ですが、ありとあらゆることを考えなければならないことがわかります。この事実気づくことが機械学習の「社会実装」に直結します。

恐らく皆さんが考えたこと

- オリンピック開催国は、インバウンド需要の増加により売上が大幅増になるはず
- 過去のオリンピック、あるいはそれに準ずるイベント時のデータで学習しているモデルを構築するのがベター
- 一方、オリンピック、あるいはそれに準ずるイベント時のデータを持っていないため、事実上それは不可能
- 通常通りにモデルが予測をすると、売上の大幅増を過小評価してしまう可能性がある。どうするべきか？
- 需要の大幅増が見込まれる場合、何らかの特徴量として投入できるモデルに改良するべきではないか？
- そこまでしなくても対応できる手段は何かないか？

我々が立ち向かわなければならないのは、まさにこのような問題の数々！