

プラント保安分野 AI 信頼性評価ガイドライン

2020 年 11 月

石油コンビナート等災害防止3省連絡会議
(経済産業省、総務省消防庁、厚生労働省)

エグゼクティブサマリー（要約）

<背景>

石油・化学プラントは設備の高経年化や保安人材の不足に直面しており、保安体制の持続性が低下していく懸念がある。

一方で、IoT（Internet of Things）、ドローン、AI（人工知能）等の新技術の実用性が高まっている。これらの新技術を適切に活用しプラントを安定的に運転することで、保安力を維持・向上できるだけでなく、製品の品質の安定化、コスト面での効率化、工期・納期の遵守などを実現できる。

特にAIについては、近年の機械学習の技術発展を背景に、プラント保安分野でもAIを組み込んだ保安システムが構築されつつある。例えば、大量のセンサデータの関係性から微細な異常の予兆を検出するAIや、蒸留装置等の生産性を高める運転の最適化AI等、保安力と生産性を飛躍的に向上させるAIの実証が進んでいる。

<課題>

実証から実装に進むためには、AIが期待通りの品質を果たすこと（＝信頼性）について適切な検証を行う必要がある。

しかし、これまでAIについては信頼性評価の体系が整備されておらず、特に安全性が重要視されるプラント保安分野において導入が進まない大きな原因の1つとなっている。

参考：プラント保安分野におけるAIの信頼性評価の必要性に関する事業者のコメント

- AIの不良が原因でプラントを止めると大損失（人的被害、経済的被害）が生じる。そのため社内の関係部署の了解を得るためには高度な信頼性評価が必要で、現状は難易度が高い。（プラントオーナー）
- 信頼性の評価が十分できていないので、重要な設備の管理をAIに任せることはできず、重要でない設備のみAIに任せようかと思う。一方で、重要でない設備はそもそも保全する必要性が低く、検査自体が不要である可能性がある。（プラントオーナー）
- AIベンダーである我々は、顧客側（プラントオーナー）に当社のAIの信頼性を理解していただくのに苦労している。（AIベンダー）

近年、AIの信頼性を含む品質確保や評価については、日本で先進的な検討が行われている。産業技術総合研究所(2020)「機械学習品質マネジメントガイドライン」やAIプロダクト品質保証コンソーシアム(2020)「AIプロダクト品質保証ガイドライン」において信頼性評価の体系や留意点がとりまとめられている。しかし、これらのガイドラインはEC（電子商取引）や自動走行などを含めた分野横断的な内容であり、プラント保安分野での信頼性評価について具体的な検討はされていない。

したがって、プラント保安分野で信頼性の高い AI の実装を進めるためには、分野横断的な AI の信頼性評価の方法論を、プラント保安分野において解釈して適用する方法について整理を行う必要がある。

＜ガイドラインの目的及び活用方法＞

本ガイドラインは、上記の課題意識のもと、プラント保安分野に特化して AI の信頼性（プラントの安全性や生産性向上のために期待される品質を果たすこと）を適切に管理する方法を示す。

本ガイドラインを活用することで、プラントオーナー企業は、AI の高い信頼性を実現し、保安力や生産性を向上させることができる。また、AI の信頼性について社内外へ説明責任を果たしやすくなる。更に、当該システムをベンダー企業とともに開発する際に要求事項を適切に伝達し、その達成状況を円滑に確認できる。

AI を開発し納品するベンダー企業は、本ガイドラインによって十分な信頼性を持つ AI を構築するための方法が明確になるため、プラントオーナー企業に対して AI の信頼性を説明しやすくなる。また、オーナー企業との AI 開発プロセスで要求事項を適切に設定できるようになることも期待される。

＜信頼性評価の構造・流れ＞

基本的な考え方

- 本ガイドラインは、産業技術総合研究所(2020)「機械学習品質マネジメントガイドライン 第1版」と同様の方法で AI の信頼性評価を行い、必要な信頼性を確保する方策を示す。その上で、プラント保安分野の AI の実用例に基づいて、プラント保安分野への具体的な適用方法を示す。
- なお、本ガイドラインは AI の中でも、近年に実用化が進んでいる機械学習の信頼性評価を対象としている。

機械学習の品質を3階層に分け（上位から「利用時品質」→「外部品質」→「内部品質」）、それらの達成を通じて機械学習利用システムの信頼性を管理する（→2）。

「利用時品質」：機械学習要素を含むシステム全体が実現したいこと（→2.1.1）。

「外部品質」：「利用時品質」を満たすために機械学習要素が満たすべきこと（→2.1.2）。「リスク回避性」「パフォーマンス」の2種類に分類する。

- リスク回避性：安全性を追求するタイプの品質。機械学習要素の誤判断によって悪影響（人的被害・経済的被害）を及ぼすリスクを回避・低減することを目的とする（→2.1.3(1)）。
- パフォーマンス：生産性を追求するタイプの品質。プラントの運転・点検を効率的に行うことを目的とする（→2.1.3(2)）。

「内部品質」：「外部品質」を満たすために機械学習要素の設計・開発・運用等で満たすべきこと。次の3種類について、8つの内部品質を定める（→2.1.4）。

- 開発に用いるデータの適切性（データの量、種類など）
- 開発したモデルの適切性（テストの際の機械学習の精度など）
- 実装・運用方法の適切性（機械学習の精度の維持の方法など）

信頼性評価の基本的な手順は次の通り（図1）。

- ① 「利用時品質」を設定する（→2.2.1）。
- ② 「利用時品質」実現のために必要な「外部品質」を設定し、「リスク回避性」「パフォーマンス」の分類ごとに「外部品質」の達成レベル（要求の強さ）を決める。ガイドラインに記載の手順に従って、リスク回避性はレベル0、0.1、0.2、1の4段階、パフォーマンスはレベル0、1、2の3段階から数値で設定する。（→2.2.2、2.2.3）。
- ③ 設定した「外部品質」のレベルの高低に応じて、要求される「内部品質」のレベル（レベル1、2、3）が決まる。（例えば、学習データの種類・量に関する内部品質は、レベルが高いほど多種・多量のデータを確実に集めることが求められる）。それによって機械学習要素の開発等を行う（→2.2.4、2.2.5）。

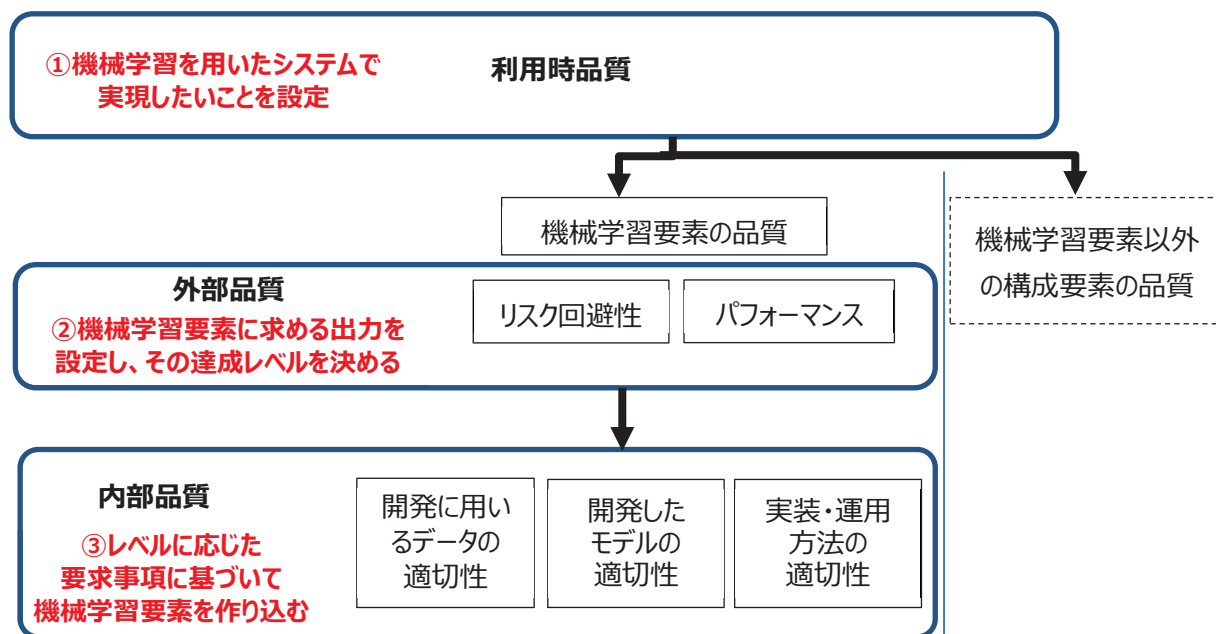


図1 信頼性評価の手順

プラント保安分野のユースケースの一例として本ガイドラインでとりあげている（→ 3.3.4）「異常予兆検知・診断」（多数のセンサデータを機械学習要素が解析して数十分～数日後に顕在化する運転中の異常の兆候を検知する機械学習）を例にとると、以下の流れで信頼性を確保することとなる（図 2）。

- ① 本ケースで実現したいことを言語化し、「利用時品質」を具体的に設定する。
- ② 設定した「利用時品質」に対応する機械学習要素の「外部品質」を具体化し、ガイドラインに記載の手順に従って「リスク回避性」「パフォーマンス」の達成レベルを決める。図 2 の例では、仮にリスク回避性はレベル 0.1、パフォーマンスはレベル 1 としている。
- ③ リスク回避性のレベル 0.1、パフォーマンスのレベル 1 には、「内部品質」のレベル 1 が該当する。レベル 1 の要求事項と、これを実行する上での留意点を確認し、機械学習要素の開発を行う。

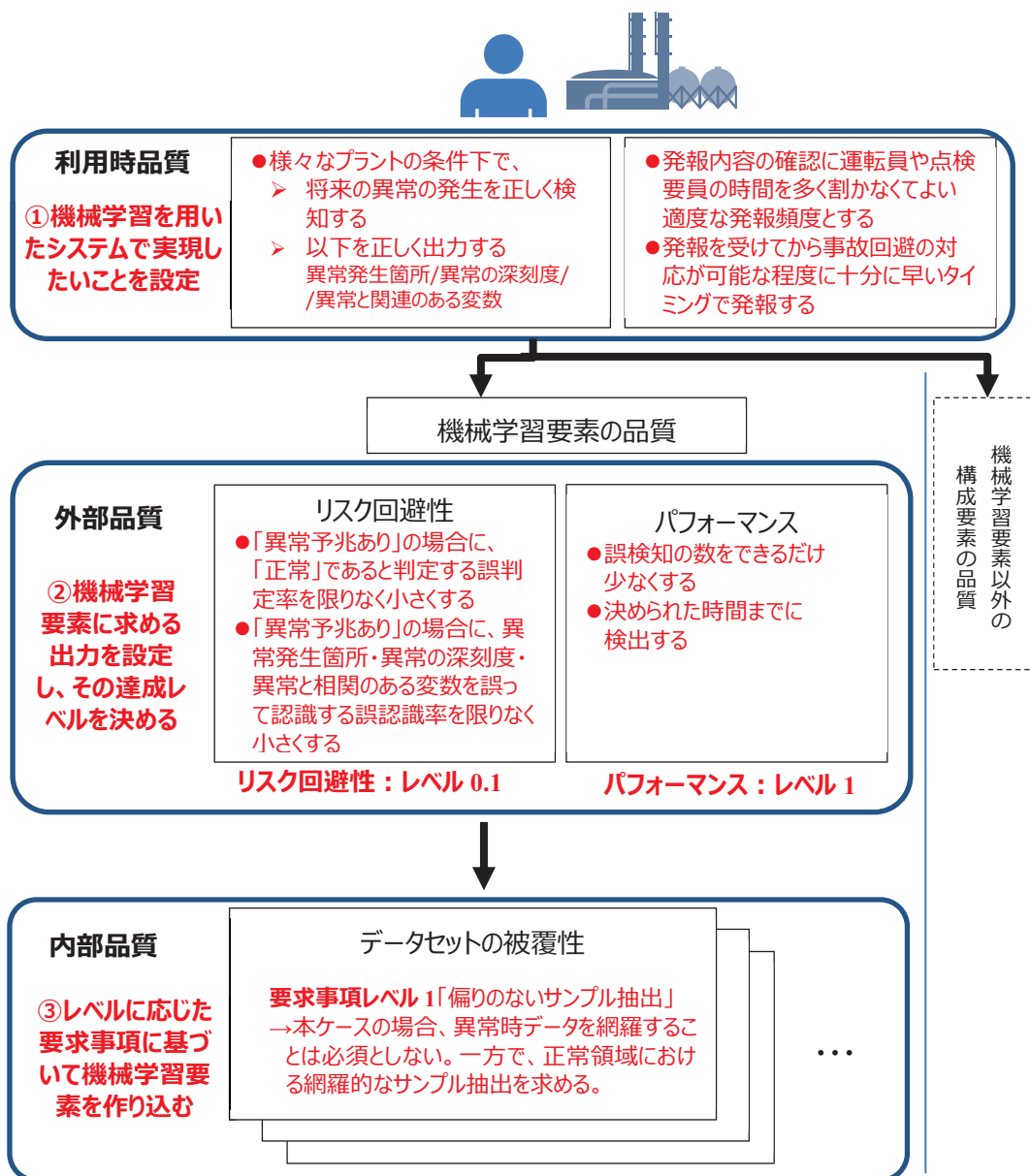


図 2 信頼性評価の流れ（「異常予兆検知・診断」を例に）

＜プラント保安分野特有の内容＞

本ガイドラインでは、上記で紹介した「異常予兆検知・診断」のほか、「運転最適化」「配管の肉厚予測」「配管の画像診断」「設備劣化診断」をユースケースとして、ガイドラインの適用方法を示している。具体的には、利用時品質・外部品質の項目の例と、内部品質の要求事項を実行する上での留意点を示している(→3)。

自社で開発している機械学習要素に近いユースケースがある場合は、ガイドラインの例を参考にすることを想定している。近いユースケースがない場合でも、信頼性評価の構

造・流れはそのまま適用できる。自社のケースに応じて利用時品質・外部品質の項目やレベルを検討することが期待される。

さらに、プラントシステム担当者、品質保証担当者、保全員等の担当者ごとにガイドラインをどのように活用すれば良いのかについて、具体的な流れに即して示している（→ 4）

加えて、ガイドラインに記載している実施事項をもれなく実施するために、個々の実施事項の検討結果を記録するフォーマットも公開している¹。フォーマットに従って実施事項を記録することで、ガイドラインを適用したエビデンス（証拠）を残すことができ、社内等への説明に活用することができる。

¹https://www.fdma.go.jp/relocation/neuter/topics/fieldList4_16/jisyuhoan_shiryo.html

目次

エグゼクティブサマリー（要約）

1. はじめに	1
1.1 ガイドラインの目的及び活用方法.....	1
1.1.1 ガイドライン策定の目的.....	1
1.1.2 本ガイドラインの活用効果.....	2
1.1.3 本ガイドラインの策定経緯.....	3
1.2 他のガイドラインとの関係.....	4
1.2.1 AI の品質に係る既存のガイドライン.....	4
1.2.2 本ガイドラインの位置づけ.....	4
1.3 適用範囲.....	6
1.4 本ガイドラインの構成及び読み方.....	6
1.5 用語.....	9
1.6 （参考）「プラントにおける先進的 AI 事例集」について.....	14
2. プラント保安分野における機械学習の信頼性評価の構造	16
2.1 信頼性評価の 3 品質.....	17
2.1.1 利用時品質.....	17
2.1.2 外部品質.....	17
2.1.3 利用時品質・外部品質の軸.....	18
2.1.4 内部品質.....	20
2.2 信頼性評価の方法と要求事項.....	27
2.2.1 利用時品質の設定.....	27
2.2.2 外部品質の設定.....	28
2.2.3 外部品質のレベル設定.....	29
2.2.4 内部品質のレベルの確認.....	37
2.2.5 内部品質の要求事項の確認および実行.....	37
3. プラント保安分野における機械学習のユースケース	40
3.1 本ガイドラインにおけるユースケースの位置づけ.....	40
3.2 対象とするユースケース.....	40
3.3 ユースケースに基づく信頼性評価の具体的適用.....	41
3.3.1 配管の肉厚予測.....	43
3.3.2 配管の画像診断.....	50
3.3.3 設備劣化診断.....	58
3.3.4 異常予兆検知・診断.....	65
3.3.5 運転最適化.....	73
4. ガイドラインの活用の流れ	82

4.1 ガイドラインの活用の主体.....	82
4.2 ガイドラインの適用の流れ.....	86
4.2.1 担当別・フェーズ別の実施項目	86
4.2.2 実施項目・実施内容の全体像	90
4.2.3 フェーズ別の実施項目・実施内容	94

附録 内部品質確保のための「プラント保安分野での観点」チェックリスト	
別紙 プラントにおける AI の信頼性評価に関する検討会 委員等名簿	

図目次

図 1-1 本ガイドラインと産業技術総合研究所ガイドラインの関係.....	5
図 1-2 プラント保安分野における AI 導入の課題の分類と「プラントにおける先進的 AI 事例集」の位置づけ	15
図 2-1 機械学習利用システムの階層的な品質確保	16
図 2-2 内部品質の 8 つの軸及びその関係性	22
図 2-3 信頼性評価における 3 段階の品質の関係	27
図 2-4 利用時品質の設定イメージ（異常予兆検知・診断システムの例）	28
図 2-5 外部品質の設定イメージ（異常予兆検知・診断システムの例）	29
図 2-6 リスク回避性のレベル(AISL)の設定フロー	30
図 2-7 機械学習要素の出力・判断に対する人間の関与.....	35
図 2-8 「機械学習品質マネジメントガイドライン 第1版」と 本ガイドラインの AIPL の考え方	37
図 2-9 外部品質の AISL/AIPL と内部品質の要求レベルの対応.....	37
図 3-1 プラントにおける事故等の進展とユースケースの関係.....	41
図 3-2 ユースケース「配管の肉厚予測」の導入イメージ（例）	44
図 3-3 ユースケース「配管の肉厚予測」における 機械学習要素と他のシステムの関係（例）	45
図 3-4 階層的な品質確保における利用時品質・外部品質項目設定の位置づけ	46
図 3-5 階層的な品質確保における「ユースケース固有の観点」の位置づけ	49
図 3-6 ユースケース「配管の画像診断」の導入イメージ（例）	52
図 3-7 ユースケース「配管の画像診断」における 機械学習要素と他のシステムの関係（例）	53
図 3-8 階層的な品質確保における利用時品質・外部品質項目設定の位置づけ	54
図 3-9 階層的な品質確保における「ユースケース固有の観点」の位置づけ	57
図 3-10 ユースケース「設備劣化診断」の導入イメージ（例）	59
図 3-11 ユースケース「設備劣化診断」における 機械学習要素と他のシステムの関係（例）	60
図 3-12 階層的な品質確保における利用時品質・外部品質項目設定の位置づけ	61
図 3-13 階層的な品質確保における「ユースケース固有の観点」の位置づけ	64
図 3-14 ユースケース「異常予兆検知・診断」の導入イメージ（例）	66
図 3-15 ユースケース「異常予兆検知・診断」における 機械学習要素と他のシステムの関係（例）	67
図 3-16 階層的な品質確保における利用時品質・外部品質項目設定の位置づけ	68
図 3-17 階層的な品質確保における「ユースケース固有の観点」の位置づけ	72
図 3-18 ユースケース「運転最適化」の導入イメージ（例）	75
図 3-19 ユースケース「運転最適化」における 機械学習要素と他のシステムの関係（例）	76
図 3-20 階層的な品質確保における利用時品質・外部品質項目設定の位置づけ	78
図 3-21 階層的な品質確保における「ユースケース固有の観点」の位置づけ	81
図 4-1 フェーズ間で反復のあるシステム構築・運用のスタイル.....	89

図 4-2 ステップ 1「システムの目的の設定」の実施イメージ.....	96
図 4-3 ステップ 2「システムの機能要件・利用時品質の設定」の実施イメージ.....	97
図 4-4 ステップ 3「外部品質の設定」の実施イメージ.....	99
図 4-5 ステップ 4「外部品質のレベルに係る機械学習利用システムと 安全関連系全体の整理」の実施イメージ.....	100
図 4-6 ステップ 5「外部品質のレベル設定」の実施イメージ.....	101
図 4-7 ステップ 6「内部品質のレベル設定」の実施イメージ.....	103
図 4-8 ステップ 7「機械学習要素の設計・開発」の実施イメージ.....	104
図 4-9 ステップ 8「機械学習要素以外の構成要素（外部安全機構等）の開発」の実施イメージ.....	105
図 4-10 ステップ 9「テスト（外部品質の評価）」の実施イメージ.....	107
図 4-11 ステップ 10「検収」の実施イメージ.....	108
図 4-12 ステップ 11「利用時品質の確認」の実施イメージ.....	110
図 4-13 ステップ 12「外部品質の確認」の実施イメージ.....	111
図 4-14 ステップ 13「内部品質の確認」の実施イメージ.....	112
図 4-15 ステップ 14「システムの更新」の実施イメージ.....	113

目 次

表 1-1	AI の「信頼性」に関するプラント保安に関わる事業者の問題意識	2
表 1-2	関心事項に応じたガイドラインの参照箇所	8
表 2-1	「リスク回避性」の簡易評価の基準（AISL 表）	33
表 2-2	本ガイドラインにおける強度基準と既存の基準・分類との関係	34
表 2-3	「パフォーマンス」のレベル設定・評価基準（AIPL）	36
表 3-1	ユースケースの概要	40
表 3-2	「配管の肉厚予測」の機械学習要素の構成（例）	45
表 3-3	ユースケース「配管の肉厚予測」の利用時品質・外部品質の項目例	46
表 3-4	「配管の肉厚予測」に係る「ユースケース固有の観点」	48
表 3-5	「配管の画像診断」の機械学習要素の構成（例）	53
表 3-6	ユースケース「配管の画像診断」の利用時品質・外部品質の項目例	54
表 3-7	「配管の画像診断」に係る「ユースケース固有の観点」	56
表 3-8	「設備劣化診断」の機械学習要素の構成（例）	60
表 3-9	ユースケース「設備劣化診断」の利用時品質・外部品質の項目例	61
表 3-10	「設備劣化診断」に係る「ユースケース固有の観点」	63
表 3-11	「異常予兆検知・診断」の機械学習要素の構成（例）	67
表 3-12	ユースケース「異常予兆検知・診断」の利用時品質・外部品質の項目例	68
表 3-13	「異常予兆検知・診断」に係る「ユースケース固有の観点」	71
表 3-14	「運転最適化」の機械学習要素の構成（例）	77
表 3-15	ユースケース「運転最適化」の利用時品質・外部品質の項目例	77
表 3-16	「運転最適化」に係る「ユースケース固有の観点」	79
表 4-1	機械学習利用システムの開発プロジェクトの担当	83
表 4-2	機械学習利用システムの開発プロジェクトにおける役割別の所属企業の例	84
表 4-3	各担当のガイドラインの活用場面と確認すべき品質	85
表 4-4	機械学習利用システム構築・運用のフェーズ・担当別の実施項目	87
表 4-5	機械学習利用システムの構築・運用におけるフェーズ別の実施項目	88
表 4-6	プラント保安分野における機械学習の信頼性評価 実施項目・内容リスト	91
表 4-7	PoC 終了段階で目処をつけることが望ましい項目	94
表 4-8	フェーズ「要求・要件定義」の実施事項	95
表 4-9	フェーズ「設計」の実施事項	98
表 4-10	フェーズ「実装」の実施事項	102
表 4-11	フェーズ「テスト・検収」の実施事項	106
表 4-12	フェーズ「運用」の実施事項	109

1. はじめに

1.1 ガイドラインの目的及び活用方法

1.1.1 ガイドライン策定の目的

AI の産業への応用の可能性と障害について、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1版」では次のように記している。

「人工知能 (AI)、とりわけ機械学習技術は、製造業、自動運転、ロボット、ヘルスケア、金融、小売などの幅広い応用分野で有効性が確認され、社会実装が本格化する兆しを見せている。その一方、AI を利用した製品・サービスの品質を測定し説明する技術の不足に起因し、万が一の事故の際に原因が特定できず、また投資に見合う AI システムの優位性を説明できず、結果として、社会的な受容性を得るための制度設計の遅れや、AI 開発ビジネス拡大への大きな障害となっている。」²

特にプラント保安分野では、「①AI が支援・代替する機能が保安（安全を守る防護機能）である」、「②多様なステークホルダーへの説明責任を負っている」という特徴があるため、AI の信頼性を説明する技術の不足の課題が顕著である。

①については、保安機能はこれまで人間や既存のシステム・ソフトウェアが担っており、適切に動作して必要な機能を発揮するための信頼性を確保・評価するための体系³が蓄積されている。それによって、保安機能は、公共や労働者の安全を確保してきた。この保安機能に、新たに AI を保安機能の支援や代替として組み込むためには、AI の動作の信頼性を確保・評価するための確立した方法論が求められることとなる。こうした方法論がなければ、安全確保の根拠が不明瞭なまま AI が保安機能に組み込まれ、安全が損なわれることになってしまう。

②については、プラントオーナー企業は、まず、社内で環境安全部門や経営層を含めて AI の蓋然的な効果や安全面でのリスクに関して適切な認識を持ち、AI の導入に合意する必要がある。加えて、社外の地域住民や規制当局等に保安活動の説明を行う必要がある。これら際には適正なプロセスで AI の開発・実装・運用等を行っているエビデンス（証拠）を示し、合意形成を行うことが求められる。

プラント保安に関わる事業者が抱えている、AI の「信頼性」に関する問題意識を表 1-1 に示す。

² 産業技術総合研究所(2020)「機械学習品質マネジメントガイドライン 第1版」

³ 法令で定める保安確保手段や、機能安全やソフトウェア品質に係る国際規格等。

表 1-1 AI の「信頼性」に関するプラント保安に関わる事業者の問題意識⁴

- 【①関連】信頼性の評価が十分できていないので、重要な設備の管理を AI に任せすることはできず、重要でない設備のみ AI に任せようかと思う。一方で、重要でない設備はそもそも保全する必要性が低く、検査自体が不要である可能性がある。（プラントオーナー）
- 【①②関連】AI の不良が原因でプラントを止めると大損失（人的被害、経済的被害）が生じる。そのため社内の関係部署の了解を得るためには高度な信頼性評価が必要で、現状は難易度が高い。（プラントオーナー）
- 【②関連】スーパー認定では、保安の手段の妥当性を自分で当局等に説明する必要がある。そのため、KHKS⁵から外れた保安検査をする場合、社内で「AI を使ってよい」と決裁が取れるまでに非常に長い時間がかかる。（プラントオーナー）
- 【②関連】AI ベンダーである我々は、顧客側（プラントオーナー）に当社の AI の信頼性を理解していただくのに苦労している。信頼性の評価についてガイドラインができれば、顧客側がそれを参照して納得頂けることになると考えられる。（AI ベンダー）

このように、プラント保安分野の文脈では、AI の信頼性を評価する体系化された方法論が特に必要とされている（＝本ガイドラインの課題）。この課題を解決するため、本ガイドラインでは、AI の中で実用化が進みつつある「機械学習」の信頼性の評価について、プラント保安分野に特化して考え方を示すことを目的としている。

1.1.2 本ガイドラインの活用効果

本ガイドラインは、主な読者として、機械学習技術を利用したシステム（以下、「機械学習利用システム」という）をプラントに導入するプラントオーナー企業と、機械学習利用システムを開発し納品するベンダー企業を想定している。それぞれの立場において想定される活用効果を以下に示す。

なお、本ガイドライン 4 章では、機械学習利用システムの構築・運用プロセスに沿って、関係するステークホルダー別にガイドラインの活用場面・参照箇所を詳述している。

(1) プラントオーナー企業の活用効果

本ガイドラインが提示する信頼性評価の方法に則ることで、保安システム内の機械学習技術を用いて構築したソフトウェア要素（以下「機械学習要素」という）に対して、適切な安全性・生産性の要求水準を設定し、その実現のために特定の要求事項を満たすことになる。このプロセスを経ることで、機械学習要素の高い信頼性を実現し、保安力や生産性を向上さ

⁴ ガイドライン策定にあたり実施したインタビュー調査による。

⁵ 高圧ガス保安協会が策定した保安検査基準。「保安検査の方法を定める告示」（平成 17 年 3 月 30 日経済産業省告示第 84 号）において、保安検査の方法として指定されている。ただし、「特定認定事業者（スーパー認定事業所）」の場合、損傷、変形及び異常の発生状況等を確認するために十分な方法であること等を事業者が判断すれば、この保安検査基準によらず事業者が自由に設定した方法で検査することが認められている。

せることができる。

さらに、本書に則った機械学習要素の構築・運用を行ったエビデンス（証拠）を示すことで、安全性等について社内外へ説明責任を果たしやすくなる。また、保安システムをベンダー企業とともに開発する際に、要求事項を適切に伝達し、その達成状況を円滑に確認することができる。

(2) ベンダー企業の活用効果

ベンダー企業は、プラントオーナー企業との機械学習の開発・運用に関する共通言語として本ガイドラインを用いることが出来る。具体的には、プラントオーナー企業が機械学習要素に求める多種多様な要求水準を、本ガイドラインが設定しているレベルに変換して受け取ることができる。更に、機械学習要素の開発に際しては、レベルに対応して本書が設けている要求事項を実施すればよく、実施事項の妥当性をプラントオーナー企業に説明する煩雑さを省略することができる。

これにより、プラントオーナー企業から機械学習要素の信頼性について理解を得やすくなる。また、適切な信頼性評価を行っていないサービスとの差別化を行うことができる。

1.1.3 本ガイドラインの策定経緯

2020 年 4 月から 11 月にプラント保安分野の事業者、安全や人工知能の専門家等による「プラントにおける AI 信頼性評価に関する検討会⁶」を開催した。検討会ではプラント保安分野における機械学習活用の状況や、それを踏まえた信頼性評価の方法等について検討を行い、本ガイドラインはその結果を踏まえて策定するものである。

今後、本ガイドラインの活用状況や、プラント保安分野の技術の進展、AI およびその品質確保・評価技術の進展等を踏まえて、ガイドラインを継続的に見直していく。

⁶ 検討会の委員等の構成は、本ガイドラインの別紙を参照。

1.2 他のガイドラインとの関係

1.2.1 AI の品質に係る既存のガイドライン

機械学習の信頼性を含む品質確保や評価については、世界的にも確立した体系が存在していないが、日本において先進的な検討が行われている。本ガイドラインの策定にあたっては、主に以下のガイドラインを参考としている。

(1) 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1版」

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1版」は、機械学習利用システムについて、「サービス提供で求められる品質要求を充足するための必要な取り組みや検査項目を体系的にまとめたもの」⁷である。機械学習利用システムの品質を「利用時品質」「外部品質」「内部品質」の3段階⁸に分け、「機械学習要素の『内部品質』の向上により『外部品質』を必要なレベルで達成し、最終的な製品の『利用時品質』を実現する」⁷ものと整理している。機械学習要素に対する品質要求（外部品質）を「リスク回避性」「AIパフォーマンス」「公平性」の3つの軸に区分⁸し、それぞれについて要求の強さに応じたレベルを設定している。また、品質要求のレベルに応じた機械学習要素を開発する際に実施すべき、開発プロセスやデータに関する要求事項を整理している。

「機械学習品質マネジメントガイドライン 第1版」は、特定の産業を想定しない汎用的な内容であり、今後、特定分野への応用にあたっては、産業分野別のリファレンスの策定を求める予定になっている。

(2) AI プロダクト品質保証コンソーシアム「AI プロダクト品質保証ガイドライン 2020.08 版」

「AI プロダクト品質保証ガイドライン 2020.08 版」は、民間企業や大学、研究機関等の個人・組織からなる「AI プロダクト品質保証コンソーシアム」により策定された、「AI プロダクトの品質保証に対する共通の指針」⁹である。AI プロダクトの品質保証の枠組みとして、Data Integrity、Model Robustness、System Quality、Process Agility、Customer Expectation の5つの軸を提示し、それぞれの軸で考慮すべき事項をチェックリスト形式で整理している。また、品質保証に有用な技術カタログの整理や、ドメイン個別のガイドラインの例示を行っている。例示されているドメインには、コンテンツ生成系システム、スマートスピーカー、産業用プロセス、自動運転、AI-OCR の5つがある。

1.2.2 本ガイドラインの位置づけ

本ガイドラインは、「機械学習品質マネジメントガイドライン 第1版」の体系を用いて

⁷ 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を公開」、

https://www.aist.go.jp/aist_j/press_release/pr2020/pr20200630_2/pr20200630_2.html、2020年9月18日アクセス

⁸ 詳細は2章を参照。

⁹ AI プロダクト品質保証コンソーシアム「AI プロダクト品質保証ガイドライン 2020.08 版」

おり、プラント保安分野への適用方法を示したリファレンスとして位置づけられる。図 1-1 に「機械学習品質マネジメントガイドライン 第1版」との関係を示す。

本ガイドラインは、「機械学習品質マネジメントガイドライン 第1版」の信頼性評価の構造を用いている。具体的には、「利用時品質」「外部品質」「内部品質」の3段階や、外部品質の2つの軸およびレベル、内部品質の8つの軸および要求事項で同様の構造をしている（図 1-1 の左半分）。

更に、この信頼性評価の構造に、プラント保安分野での実例に基づいて肉付けを行っている。具体的には、利用時品質・外部品質の具体的な項目例や、システム全体の安全機能を踏まえたレベルの設定手順、要求事項を実現するための具体的な観点を示している（図 1-1 の右半分）。

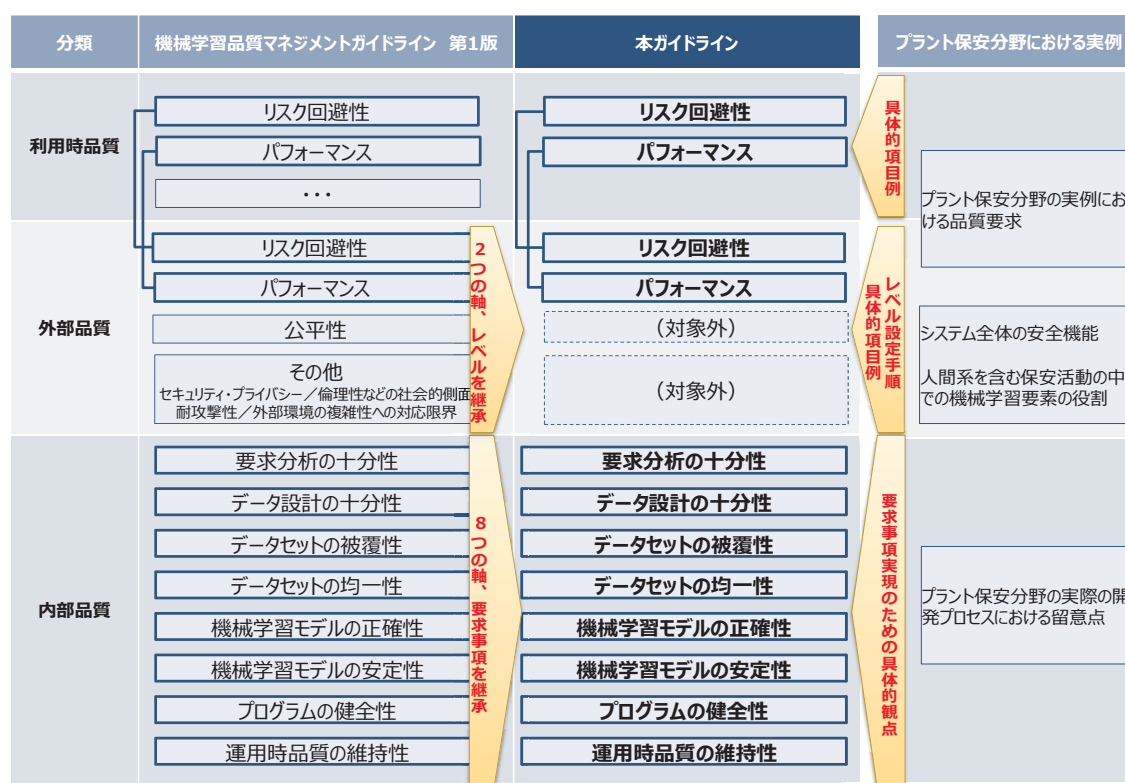


図 1-1 本ガイドラインと産業技術総合研究所ガイドラインの関係

また、「AI プロダクト品質保証ガイドライン 2020.08 版」の7章で、について、プラント制御に機械学習要素を適用することを題材とした「産業用プロセス」ドメインの品質保証が扱われている。ここで挙げられている留意事項について、「要求事項実現のための具体的な観点」として追加している（2.1.4 に詳述）。

1.3 適用範囲

本ガイドラインは、機械学習要素の信頼性を扱い¹⁰、プラント保安分野に機械学習を適用する場合に、信頼性の観点から、機能安全など従来からのプラント保安の信頼性確保の取組に追加して検討すべき事項をまとめたものである¹¹。

例えば、本ガイドラインで扱う「リスク」とは、機械学習要素の誤判断に起因する危険・災害である。ハードウェアの信頼性・危険側故障は本ガイドラインでは取り扱っておらず、処理を実行するハードウェアの故障に起因する危険・災害は対象になっていない。また、機械学習要素以外のソフトウェアの信頼性や、プラントの保安を担う人員の業務（手順、マニュアル等）も本ガイドラインでは取り扱っていない。

すなわち、本ガイドラインに基づく検討をもって、機械学習要素を含むシステム全体、ひいてはプラント全体の安全や性能が実現されるものではない。

加えて、法令上・倫理上の問題、第三者のプライバシー、社会的受容の確保、サイバーセキュリティといった、新技術・データ活用において一般的に課題とされている事項についても、本ガイドラインでは対象とせず、別途検討すべき事項とする。

なお、本ガイドラインは法令の規定を緩和したりその解釈を示したりするものではなく、法定検査に機械学習要素を活用する場合は法定義務を遵守する必要がある。

導入する機械学習要素の構築プロセスが本ガイドラインに適合しているかについてプラントオーナー企業が規制当局に説明する際には、プラントオーナー企業が説明責任を負う。

1.4 本ガイドラインの構成及び読み方

本ガイドラインは以下の構成となっている。

2章では、「プラント保安分野における機械学習の信頼性評価の構造」として、本ガイドラインの階層的な信頼性評価の方法を示す。読者は、適切な信頼性評価のためには、本章に則った機械学習要素の構築・運用が求められる。

2.1節では「信頼性評価の3品質」として、信頼性評価の3階層（「利用時品質」「外部品質」「内部品質」）のそれぞれの意味を、プラント保安分野における解釈も含めて示す。

2.2節では「信頼性評価の方法と要求事項」として、3段階の品質に基づいて信頼性評価を実施する方法と評価基準を示す。

3章では「プラント保安分野における機械学習のユースケース」として、プラント保安分野の典型的なユースケースに基づいて具体的な信頼性評価の例を示す。本章は、読者が2章の内容を実践する際に参照することを想定しており、読者が構築・運用する機械学習要素に

¹⁰ 機能安全規格では、「JIS C 0508-3:2014 電気・電子・プログラマブル電子安全関連系の機能安全ー第3部：ソフトウェア要求事項」の対象に該当する。

¹¹ プロセス産業分野の安全計装システム（機能安全）についての国際規格として、IEC 61511-1(JIS C 0511-1)がある。ここでは、ソフトウェアを、「3.2.75.1 FPL(fixed program language: 固定プログラム言語)」、「3.2.75.2 LVL(Limited variability language: 制約可変言語)」、「3.2.75.3 FVL(Full variability language: 完全可変言語)」のいずれかにより実現されるものと分類している。この中で、機械学習要素はFVLにより実装される要素として考えるべきである。FVLについては、IEC 61508-3:2010(JIS C 0508-3:2014)により取り扱うものとされているため、本ガイドラインの対象とする「プラント保安分野の機械学習要素の信頼性」は、脚注10の通り、IEC 61508-3:2010(JIS C 0508-3:2014)の対象に該当するものと考えられる。

合わせて適用することが期待される。

4章では「ガイドラインの活用の流れ」として、機械学習利用システムの構築・運用の流れに従って信頼性評価の実施項目を整理し、フェーズ別・関与する主体別に、本ガイドラインの活用方法例を示す。本章は、読者が2章の内容を実践する際に参照することを想定しており、読者のプロジェクトの状況に合わせて柔軟に活用することが期待される。

加えて、附録として、3段階の品質のうち「内部品質」に対する要求事項と、要求事項をプラント保安分野で実行するための留意点を整理し、チェックリストとして提示する。

なお、ガイドラインに記載している実施事項をもれなく実施するために、個々の実施事項の検討結果を記録するフォーマットも公開している¹²。フォーマットに従って実施事項を記録することで、ガイドラインを適用したエビデンス（証拠）を残すことができ、社内外での説明等に活用することができる。

本ガイドラインは、1章から順に読むことで、信頼性評価の方法と適用例、実施項目が理解できるよう構成しているが、読者の関心に応じて特定の章・節から読むことも想定している。関心事項に応じたガイドラインの参照箇所は次の表の通りである。

¹² https://www.fdma.go.jp/relocation/neuter/topics/fieldList4_16/jisyuhoan_shiryo.html

表 1-2 関心事項に応じたガイドラインの参照箇所

関心事項	ガイドラインの参照箇所
ガイドラインのどの部分を誰(品質保証担当、プラントシステム担当、機械学習開発・設計担当等)が理解すればいいのか把握したい	4.1 節
ガイドラインを誰がどのように使うのかを把握したい	4.2 節
自社で検討している AI に近い信頼性評価の検討例を把握したい	3.1 節・3.2 節でユースケースの位置づけ・一覧を確認、ユースケースを選定 →3.3 節のうち該当するユースケースを参照 3.3.1: 配管の肉厚予測 3.3.2: 配管の画像診断 3.3.3: 設備劣化診断 3.3.4: 異常予兆検知・診断 3.3.5: 運転最適化
信頼性評価の 3 つの品質(利用時品質、外部品質、内部品質)の意味を理解したい	2.1 節
信頼性評価の 3 つの品質に基づく、具体的な信頼性評価の方法を理解したい	2.2 節
内部品質の具体的な要求事項・観点を確認したい	附録
用語の定義を確認したい	1.5 節
ガイドラインの適用状況を記録したい	実施内容記録フォーマット

1.5 用語

本ガイドラインに使用する用語を以下のように定義する。

(1) 機械学習

人工知能のうち、特に計算機システムが明示的なプログラム指示に依ること無く、データのパターンを自動的に認識した結果を用いて推論・判断を行うようなシステム。深層学習も機械学習の代表例である。¹³

(2) 機械学習要素

機械学習技術を応用して実装されたソフトウェアコンポーネント。¹³プラント保安分野においては、例えばプロセスデータから異常の予兆を検知するソフトウェアや、配管の画像から腐食の程度を判定するソフトウェアなどを指す。

本ガイドラインは、機械学習要素の品質のみを対象としている。

(3) 機械学習利用システム

機械学習要素をコンポーネントとして内包するシステム。¹⁴プラント保安分野においては、例えば機械学習を活用した異常予兆検知・診断システムや、機械学習による運転最適化機能を組み込んだ制御システム、ドローンで撮影した配管の画像から腐食の程度を判定し可視化するシステムなどを指す。

本ガイドラインでは、機械学習利用システムに内包される、機械学習要素以外の要素（ソフトウェア／ハードウェア）の品質は対象となっていない。

(4) 教師あり学習／教師なし学習／強化学習

機械学習の学習手法は、「教師あり学習」「教師なし学習」「強化学習」の3つに分けることができる。

「教師あり学習」は、入力データから出力データのパターンを予測／識別する学習手法である。与えられた入力データと正解となる出力データのセットを学習データとし、入力と出力の関係を学習する。教師あり学習で解く問題の代表例が「回帰」と「分類」である。本ガイドラインでは、「配管の肉厚予測（回帰）（3.3.1）」「配管の画像診断（分類）（3.3.2）」「設備劣化診断（分類）（3.3.3）」のユースケースが該当する。

「教師なし学習」は、学習データとして正解を与えずに学習を行う学習手法である。学習データを特徴によってグループ分けする「クラスタリング」などが代表例。本ガイドラインでは、「異常予兆検知・診断（3.3.4）」のユースケースが該当する。

「強化学習」は、ある環境において行動を選択することで報酬を得るという枠組みにおいて、得られる報酬が最大になるような方策を学習する学習手法である。学習データには正解

¹³ 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を公開」

https://www.aist.go.jp/aist_j/press_release/pr2020/pr20200630_2/pr20200630_2.html、2020年9月8日閲覧

¹⁴ 産業技術総合研究所(2020)「機械学習品質マネジメントガイドライン 第1版」

を含まない。本ガイドラインでは、「運転最適化（3.3.5）」のユースケースが該当する。

(5) 回帰モデル／分類モデル

教師あり学習のうち、配管の肉厚など連続値を予測する問題を回帰問題、配管の腐食有無など離散値（カテゴリ）を判別する問題を分類問題といい、それぞれの問題で用いるモデルを回帰モデル、分類モデルと呼ぶ。本ガイドラインでは、回帰モデルの例として「配管の肉厚予測（3.3.1）」、分類モデルの例として「配管の画像診断（3.3.2）」「設備劣化診断（3.3.3）」のユースケースをとりあげている。

(6) プラント

石油コンビナート地域を含む石油精製、化学工業（石油化学を含む）等の事業所。¹⁵

(7) 安全関連系

次の両方を満足するシステム。

－EUC(equipment under control: 被制御機器)を安全な状態に移行させるため、又は EUC の安全な状態を維持するために必要な、要求された安全機能を行う。

－それ自体で、又はその他の E/E/PE(electrical/electronic/programmable electronic: 電気・電子・プログラマブル電子)安全関連系及び他リスク軽減措置によって、要求される安全機能に対して必要な安全度を達成する。¹⁶

ここで、EUC とは、本ガイドラインでは主に製造や保全に用いられる機器、機械類、装置、プラントなどを指す。¹⁶また、E/E/PE 系とは、動力源供給装置、入力装置（センサ）、インタフェース及び他の通信経路、出力装置（アクチュエータなど）などの要素を含むシステムを指す。¹⁶

本ガイドラインでは、単に安全関連系という場合は機械学習要素が安全関連系に含まれている場合を想定する。機械学習要素の誤判断による影響が一定以上に大きい場合、安全関連系全体の SIL 評価を求めることとしている。なお、機械学習要素を含まない安全関連系を指す場合は、次項(8)「機械学習利用システムと独立の安全関連系」を用いる。

(8) 機械学習利用システムと独立の安全関連系

機械学習利用システムの入力・処理・出力とは無関係に、対象設備の安全な状態を確保している安全関連系。機械学習要素の望ましくない出力を補正する「外部安全機構」とは異なる。

本ガイドラインでは、機械学習利用システムと独立の安全関連系がある場合には、機械学習要素に対する「リスク回避性」の考慮が不要になる場合もあるとしている（2.3.2 節）。

¹⁵ 石油コンビナート等災害防止 3 省連絡会議(2019)「プラントにおけるドローンの安全な運用方法に関するガイドライン」

¹⁶ JIS C 0508-4 : 2012 電気・電子・プログラマブル電子安全関連系の機能安全－第 4 部：用語の定義及び略語

(9) SIL(safety integrity level: 安全度水準)

機能安全規格 IEC 61508 (JIS C 0508) および IEC 61511 (JIS C 0511) で定められている、安全度の値の範囲に対応する離散的水準。安全度水準 4 は最高の安全度水準であり、1 は最低である。¹⁷

安全度とは、ある E/E/PE 安全関連系が、指定した期間内に、全ての指定した条件下で、規定する安全機能を果たす確率¹⁷を指す。SIL は、安全機能に対する安全要求事項を規定するために用いられる。

本ガイドラインでは、SIL の 4 水準 (1~4) のうち、SIL2 以上を機械学習要素に割り当てることを不可としており、「SIL1」または「SIL なし」の水準の範囲で、機械学習要素の「リスク回避性」のレベルを設定している。

(10) 外部安全機構

安全性を向上させることを目的として機械学習要素と並列・直列に処理されるソフトウェアまたはハードウェアで、機械学習要素の望ましくない出力に対して監視・補正（出力の制限や上書き修正）を行い、かつ IEC 61508(JIS C 0508) 、IEC 61511(JIS C 0511)などの機能安全規格に従う既存のシステム開発プロセスの手法で十分に安全を確保できると評価できるもの。¹⁸本ガイドラインでは、外部安全機構の有無により「リスク回避性」の要求水準が変わる、すなわち、外部安全機構があれば「リスク回避性」の要求水準が下がるものとしている。

(11) 信頼性（ソフトウェア信頼性）

特定条件下で使用する時のある性能を維持する能力。¹⁹

本ガイドラインでは、機械学習要素が期待通りの品質を果たすことを念頭に、例えば機械学習要素が配管の画像から腐食を判定する判断の確からしさ等を指す。

なお、本ガイドラインは機械学習要素の信頼性を取り扱っているため、単に信頼性という場合はソフトウェア信頼性を指すものとする。

(12) リスク

危害の発生確率及びその危害の度合いの組合せ。²⁰

本ガイドラインで取り扱うリスクは、機械学習要素の誤判断に伴うリスクのみである。

¹⁷ JIS C 0508-4 : 2012 電気・電子・プログラマブル電子安全関連系の機能安全—第 4 部：用語の定義及び略語

¹⁸ 産業技術総合研究所(2020)「機械学習品質マネジメントガイドライン」を参考に本ガイドラインで定義。

¹⁹ JIS Z 8115 : 2019 ディペンダビリティ（総合信頼性）用語

²⁰ JIS Z 8051 : 2015 安全側面—規格への導入指針

(13) 安全

許容不可能なリスクがないこと。²¹

本ガイドラインでは、機械学習要素の誤判断に伴うリスクが許容可能なレベルで管理されている状態を安全としている。

(14) 微小な怪我

不休災害のうち、医師の診察を要しないかすり傷程度のもの。(いわゆる「赤チン災害」)

リスク回避性の要求レベルを設定する際、機械学習要素の誤判断に伴う影響を判定する強度基準のうちの1つ。

(15) 利用時品質

システム全体で最終的な利用者に提供すべき品質(特定の利用状況において、利用者のニーズ(性能、使い勝手など)をどれだけ実現できるか)。²²プラント保安分野では、「最終的な利用者」は機械学習利用システムの目的に応じ、運転員・保全員である場合、個別設備の安全性・生産性に責任を持つ管理者である場合、プラント全体の管理者である場合、プラントオーナー企業の経営層である場合等様々である。

本ガイドラインでは、利用時品質のうち、機械学習要素の外部品質と対応関係にある項目の実現について取り扱っており、機械学習要素以外の要素と対応する利用時品質は対象としていない。

(16) 外部品質

システムのうち機械学習で構築された構成要素が満たすことが期待される品質。²²

外部品質を達成することで、利用時品質を実現する。

本ガイドラインでは、単に外部品質という場合は機械学習要素の外部品質を指す。

(17) リスク回避性(利用時品質・外部品質の2軸のうちの1つ)

機械学習要素の望ましくない判断動作によって、その製品の運用者・利用者または第三者などに人的被害や経済損失などの悪影響を及ぼすことを回避する品質特性である。「リスク回避性」の要求を設定し、実現することが、安全分野における「リスク低減」の概念と対応する。²²利用時品質・外部品質の2種類のうちの1つ。

本ガイドラインでは、異常予兆検知・診断における異常の見逃しや、配管の画像診断における腐食の見逃し等、プラントの安全面に悪影響を及ぼすことを回避することを指す。2.1.3項に詳述。

なお、ここでいう「回避」とは、リスクマネジメント²³におけるリスク対応の種類の1つ

²¹ JIS C 0511-1 : 2019 機能安全—プロセス産業分野の安全計装システム—第1部 : フレームワーク, 定義, システム, ハードウェア及びアプリケーションプログラミングの要求事項

²² 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1版」

²³ JIS Q 0073 : 2010 リスクマネジメント—用語

である「リスク回避(活動からの撤退)」を意味するものではない。

(18) パフォーマンス（利用時品質・外部品質の2軸のうちの1つ）

機械学習要素が、生産性・効率性向上に資する判断をすることを求める利用時品質および外部品質を総称する軸のこと。具体的には、機械学習利用システムの利用者が期待する出力（「リスク回避性」に含まれる「安心・安全」を追求する出力を除く）を、長期的に平均してより高い精度・確率で出力することを指す。個別の誤判断による悪影響が大きな問題にならない出力では、個々の出力の是非よりも、平均的な性能の高さが要求されることに着目している。²⁴

本ガイドラインでは、異常予兆検知・診断における誤検知（異常がない場合に異常と発報する）の頻度や、配管の肉厚予測における安全側の誤差（処置が不要の場合に必要と判断する）等、個別の誤判断は許容されるが、長期的に平均してより高い性能を発揮することが期待されることを指す。2.1.3 項に詳述。

(19) 内部品質

機械学習による構成要素が固有に持つ品質。²⁵「外部品質」を満たすために機械学習要素の設計・開発・運用等で満たすべきことであり、次の3種類について、8つの軸を定める（それぞれの品質の説明は2.1.4に詳述）。

- ・ 開発に用いられるデータの適切性（データの量、種類など）
- ・ 開発したモデルの適切性（テストの際の機械学習の精度など）
- ・ 実装・運用方法の適切性（機械学習の精度の維持の方法など）

(20) ユースケース

プラント保安分野における機械学習利用システム活用の典型例。本ガイドラインでは、5件のユースケースに基づいて、利用時品質・外部品質の項目の例と、内部品質の要求事項を実行するための留意点を示している。3章に詳述。

(21)（プラント保安分野での／ユースケース固有の）観点

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」に記載の内部品質の要求事項を、プラント保安分野で実行するための留意点。ユースケースを問わずプラント保安分野で共通的に参照すべき留意点「プラント保安分野での観点」と、個別のユースケースの開発に特化して参照すべき留意点「ユースケース固有の観点」がある。例えば、以下のような観点がある。

- ・ 「偏りのないデータのサンプル抽出」という要求事項を実現するにあたり、実測値データの量に偏りがある場合、シミュレータの利用を検討する（プラント保安分野での

²⁴ 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1版」を参考に記載

²⁵ 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を公開」

https://www.aist.go.jp/aist_j/press_release/pr2020/pr20200630_2/pr20200630_2.html、2020年9月11日閲覧

観点)。

- ・ 「偏りのないデータのサンプル抽出」という要求事項を実現するにあたり、異常予兆検知・診断のユースケースにおいては、異常時データを網羅することは必須とせず、正常領域における網羅的なサンプル抽出を求める（ユースケース固有の観点）。

(22) PoC

「Proof of Concept（概念実証）」の略。新しいアイデアの実施可能性を確認するために、本格的な実装・システム化の前に行う検証活動を指す。

PoC は、純粋な試行的検討から本格開発の準備まで多様な意味を含む概念であり、すべての場合に信頼性評価の検討が求められるものではないが、本ガイドラインでは PoC の段階で確認しておくことが望ましい事項を 4.2.3 節に示している。

(23) レビュー

本ガイドライン 4 章において用いている、信頼性評価への関与の仕方を表す用語。本ガイドラインを用いて信頼性評価を行う主担当から依頼を受け、自身の業務所掌・専門性に基づいて主担当による検討事項を確認する。

なお、レビューを行う担当者は、必ずしもガイドラインの内容を読んで理解する必要は無く、主担当者からの依頼に応える形で信頼性評価に関与する。

1.6（参考）「プラントにおける先進的 AI 事例集」について

なお、図 1-2 に示すように、プラント保安分野における AI 導入の課題は、信頼性の確保・評価に係るものだけではなく、人材や体制、経済的利点の不明瞭さ等様々なものがある。これらの課題の全体像や、課題を乗り越えた事例、具体的成果等については、「プラントにおける先進的 AI 事例集 ～AI プロジェクトの成果実現と課題突破の実践例～」を参照されたい。

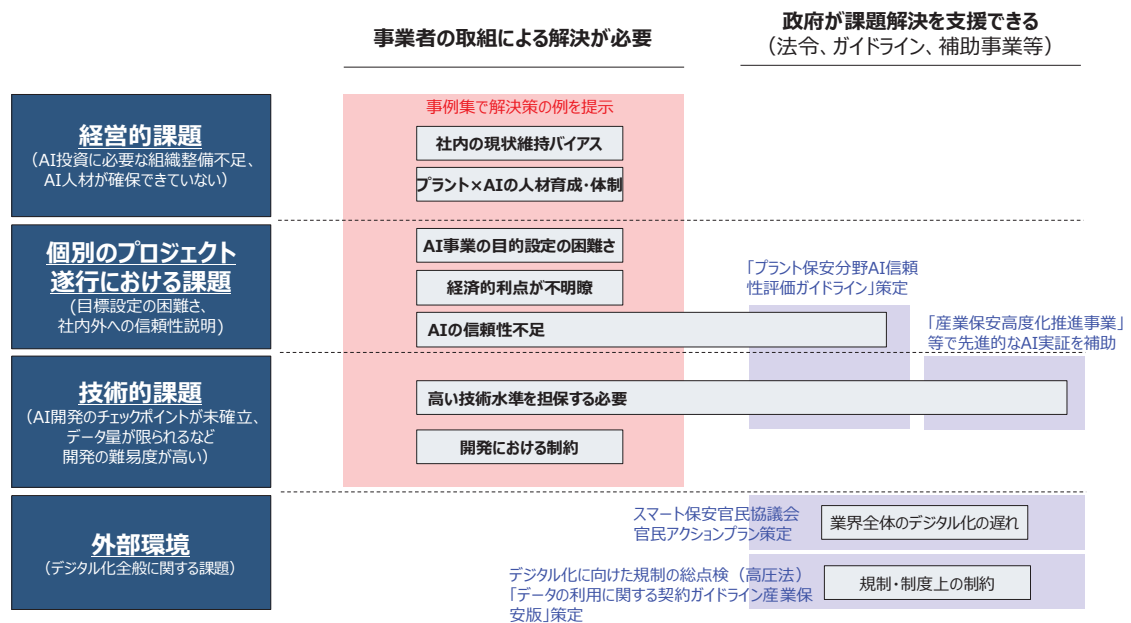


図 1-2 プラント保安分野における AI 導入の課題の分類と「プラントにおける先進的 AI 事例集」の位置づけ²⁶

²⁶ 石油コンビナート等災害防止 3 省連絡会議（経済産業省、総務省消防庁、厚生労働省）(2020)「プラントにおける先進的 AI 事例集 ～AI プロジェクトの成果実現と課題突破の実践例～」

2. プラント保安分野における機械学習の信頼性評価の構造

本ガイドラインでは、「機械学習品質マネジメントガイドライン 第1版」の階層的な品質確保の構造を用いて、「機械学習要素の『内部品質』」の向上を通じてその『外部品質』を必要となるレベルで達成し、最終的なシステムの『利用時品質』を実現する²⁷。

階層的な品質確保の手順の概略を図 2-1 に示す。①機械学習利用システムで実現したいこと（利用時品質）を設定する、②利用時品質を満たすために、機械学習要素に求める性能（外部品質）を明らかにし、求めるレベル（＝要求度合いの強さ）を設定する、③レベルに応じた要求事項に基づいて実際に機械学習要素（内部品質）を作り込む、という流れとなる。

2.1 で、「利用時品質」「外部品質」「内部品質」について説明した後で、2.2 では上記の階層的な信頼性評価の方法を詳述する。

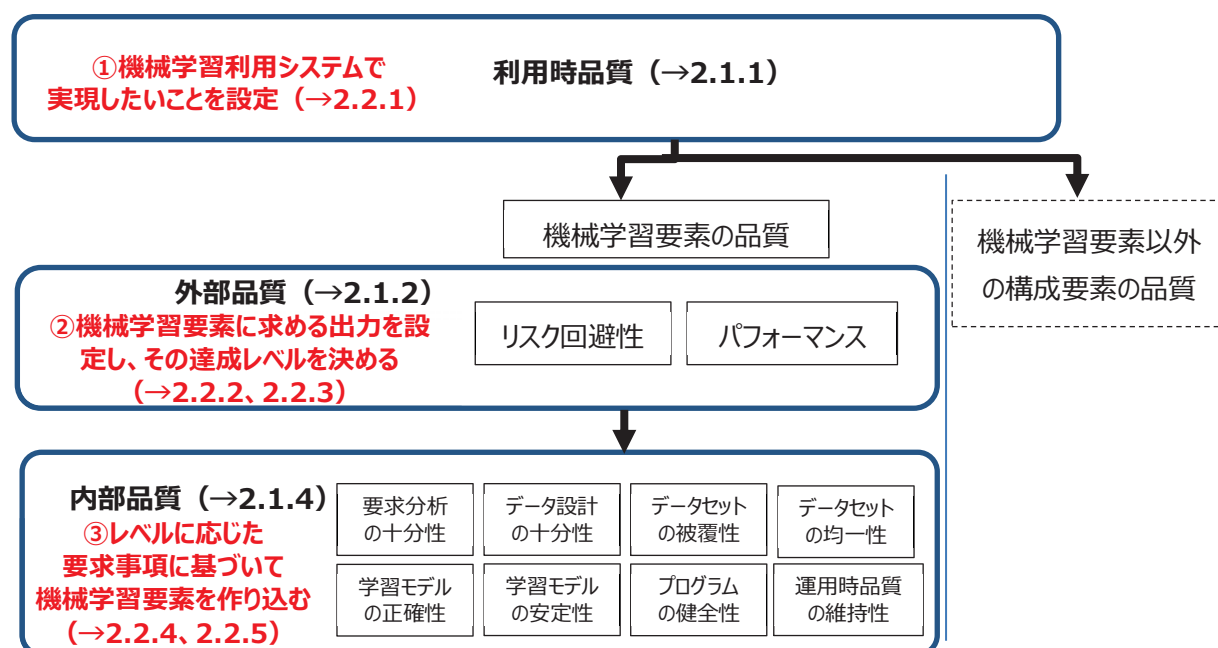


図 2-1 機械学習利用システムの階層的な品質確保

出所) 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1版」を参考に三菱総合研究所作成

なお、本ガイドラインは機械学習要素の品質のみを検討対象としているが、機械学習利用システムの品質確保においては、機械学習要素以外の構成要素（ルールベースでプログラミングされた構成要素等）の品質確保も重要である。機械学習要素以外の構成要素の品質は、既存の品質確保・評価の体系（機能安全やソフトウェア品質に係る国際規格等）に従って確保する。さらに、機械学習要素と、他の構成要素をどのように組み合わせるか（例えば、機械学習要素の出力結果と、他の構成要素の出力結果を比較し、矛盾する場合は人の判断を求め、矛盾しない場合は自動的に操作する、といったルールをどのように決めるか）によっても、機械学習利用システムとして達成される品質は変化する。本ガイドラインで対象としている機械学習要素の品質のみならず、「機械学習をシステムの中でどのように使えば機械学習利用システム全体の品質を上げることができるか」を考えることが肝要である。

²⁷ 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1版」

2.1 信頼性評価の3品質

2.1.1 利用時品質

機械学習利用システム全体で最終的な利用者に提供すべき品質²⁸が利用時品質である。これを達成するために、機械学習要素を含むシステムの構成要素を開発する。

利用時品質には、システムの利用者がシステムに期待している目的や、前提として確保されているべき安全性等が含まれており、換言すると「システムで実現したいこと」が利用時品質であるといえる。プラント保安分野においては、「異常予兆検知・診断システム」や「配管の画像診断システム」といった機械学習利用システムを特定のプラントに導入するにあたり、利用者が期待する目的や、前提として備えるべき安全性等が、利用時品質に該当する²⁹。

【利用時品質の例】

- ・ (異常予兆検知・診断システムの場合)
様々な条件下で、将来の異常の発生を正しく検知する
- ・ (配管の画像診断システム)
目視点検が必要な配管を見落とさない

ここで、「システムの利用者」は、機械学習利用システムの目的に応じて、運転員・保全員、個別設備の安全性・生産性に責任を持つ管理者、プラント全体の管理者、プラントオーナー企業の経営層等が考えられ、当該システムの利用者の立場で利用時品質が規定する。このため、利用時品質を言語化した際には、定性的な表現が中心となる。利用時品質をシステムとして達成するためには、システムの構成要素が所定の性能を発揮することが必要となる。これが次項の「外部品質」である。

2.1.2 外部品質

システムのうち機械学習を用いる構成要素が満たすことが期待される品質²⁸が外部品質である。

外部品質は、利用時品質を達成するために、機械学習要素に求められる性能として表現される。プラント保安分野においては、「異常予兆検知・診断システム」や「配管の画像診断システム」の機械学習要素（予兆の検知や腐食の判定）に求められる性能が外部品質に該当する。³⁰

²⁸ 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1版」

²⁹ 「機械学習品質マネジメントガイドライン 第1版」では、他分野における利用時品質の具体例として、以下のようなものが挙げられている。

自動運転を行う自動車の例：あらゆる運転可能な環境条件下で障害物に衝突しない安全性

株の自動取引を行うサービスの例：利益の最大化

³⁰ 「機械学習品質マネジメントガイドライン 第1版」では、他分野における外部品質の具体例として、以下のようなものが挙げられている。

自動運転を行う自動車に搭載される物体認識モジュールの例：想定される天候や時間帯などの全てにお

【外部品質の例】

- ・（異常予兆検知・診断システムの場合）
「異常予兆あり」の場合に、「正常」であると判定する誤判定率を限りなく小さくする
- ・（配管の画像診断システム）
目視点検が「要」である場合に、「不要」であると判定する誤判定率を限りなく小さくする

機械学習利用システムを、機械学習要素とその他の構成要素からなると考えると、利用時品質は「機械学習要素の外部品質」と「その他の構成要素の外部品質」により達成される。「その他の構成要素の外部品質」は、既存の品質確保・評価の体系（機能安全やソフトウェア品質に係る国際規格等）に従って確保することを前提とし、本ガイドラインでは、単に外部品質という場合は「機械学習要素の外部品質」を指すものとしている。

2.1.3 利用時品質・外部品質の軸

利用時品質・外部品質は、特徴に応じた分類を設けており、本ガイドラインでは「軸」と称している。「軸」とは、個別の機械学習利用システム及び機械学習要素について、最終的に達成したい品質の分類であり、本ガイドラインでは以下に説明する「リスク回避性」「パフォーマンス」の2軸を設定している³¹。全ての利用時品質と外部品質は、どちらかの軸に属する。

また、外部品質については、「リスク回避性」「パフォーマンス」のそれぞれの軸について要求の強さに応じたレベルを定義しており（2.2.3 参照）、外部品質のレベルに応じて、次項の「内部品質」の要求事項のレベルが決まる構造となっている。

(1) リスク回避性

機械学習要素の誤判断によって、その機械学習利用システムの運用者・利用者または第三者などに人的被害や経済損失などの「安全・安心」に関する悪影響を及ぼすことを回避³²す

いて、障害物を正しく認識する

株の自動取引を行うサービスに内包される株価予想モジュールの例：株価予測の誤差の最小化や、想定される取引結果の総和を最大化する

³¹ 「機械学習品質マネジメントガイドライン 第1版」では、機械学習要素の外部品質を「リスク回避性」「パフォーマンス」「公平性」の3つの軸としている。このうち「公平性」については、機械学習利用システムが消費者等の市民に影響を与えることを念頭に、機械学習要素に社会規範性や倫理性を機械学習に求める軸である。しかし、プラント保安分野においては、機械学習の対象はプラント設備であり市民ではないことから、本ガイドラインでは「公平性」の軸は不要であると判断した。なお、本ガイドラインでは対象としていないが、例えば従業員の画像・動画データや会話内容等をデータとして含むユースケースの場合は、「公平性」の軸での検討が必要となる可能性がある。

³² ここでいう「回避」とは、リスクマネジメントにおけるリスク対応のタイプの1つである「リスク回避（活動からの撤退）」を意味するものではない。

る利用時品質および外部品質を総称する軸である。「リスク回避性」軸に属する外部品質の要求を設定し、実現することは、安全分野における「リスク低減」と同等のことを意味する。

33

プラント保安分野における、「リスク回避性」にあたる利用時品質・外部品質の例を以下に示す³⁴。

- ・ (異常予兆検知・診断システムの場合)
利用時品質：様々なプラントの条件下で、将来の異常の発生を正しく検知する
外部品質：「異常予兆あり」の場合に、「正常」であると判定する誤判定率を限りなく小さくする
(機械学習要素が異常の兆候を見逃し、実際に異常が発生することを回避するための品質)
- ・ (配管の画像診断システムの場合)
利用時品質：目視点検が必要な配管を見落とさない
外部品質：目視点検が「要」である場合に、「不要」であると判定する誤判定率を限りなく小さくする
(機械学習要素が劣化した配管を安全と判断し、深刻な損傷や事故が発生することを回避するための品質)

異常の見逃しや、配管の画像診断における腐食の見逃し等、プラントの安全面に悪影響を及ぼすことを回避する品質を指している。このような項目に対し、悪影響を回避する要求の強さ等に応じた外部品質のレベルを設定することとなる（具体的なレベルの設定方法については、2.2.3(1)で詳述）。

(2) パフォーマンス

機械学習要素が、生産性・効率性向上に資する判断をすることを求める利用時品質および外部品質を総称する軸のこと。具体的には、機械学習利用システムの利用者が期待する出力（「リスク回避性」に含まれる「安心・安全」を追求する出力を除く）を、長期的に平均してより高い精度・確率で出力する品質を指す。個別の誤判断による悪影響が大きな問題にならない出力では、個々の出力の是非よりも、平均的な性能の高さが要求されることに着目している。³⁵

プラント保安分野における、「パフォーマンス」に属する利用時品質・外部品質の例を以下に示す。

- ・ (異常予兆検知・診断システムの場合)
利用時品質：発報内容の確認に運転員や点検要員の時間を多く割かなくてよい適度な発報頻度とする
外部品質：誤検知の頻度を一定以下にする

³³ 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1版」

³⁴ 3節では、5件のユースケースについて外部品質の項目例を記載している。

³⁵ 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1版」を参考に記載

- ・ (配管の画像診断システムの場合)

利用時品質：保全員の目視点検実施数を少なく抑える

外部品質：目視点検が「不要」である場合に、「要」であると判定する誤判定率を一定以内にす

誤検知（異常がない場合に異常と発報する）の頻度や、危険側の誤差（処置が不要の場合に必要と判断する）等、個別の誤判断は許容されるが、長期的に平均してより高い性能を発揮する（＝誤検知や危険側の誤差等を少なくする）ことが期待される品質を指している。このような項目に対し、要求の強さに応じた外部品質のレベルを設定することとなる（具体的なレベルの設定方法については、2.2.3(2)で詳述）。

なお、「リスク回避性」と「パフォーマンス」は必ずしも相反するものではなく、同一の機械学習利用システムにおいて同時に求められる場合が多い。例えば、保全に用いるシステムの場合、劣化の見逃しを避けるための「リスク回避性」を求めると同時に、点検・交換の頻度を最小化するための「パフォーマンス」も併せて求められる。それぞれの軸について求められる外部品質のレベルを設定し、より厳しいレベルに基づいて内部品質を作り込むことが必要になる。

2.1.4 内部品質

機械学習要素は、開発者の演繹的なプログラミングによって異常／正常等の判断をしないため、誤判断は様々な要因によって起こりうる（学習データセットの不十分、学習データセットへの過学習、実装環境の変化への適応不足など）。そのため、外部品質を管理するためには、プログラム（コード）の妥当性の評価だけでは不十分であり、機械学習の設計から運用までの全工程の妥当性を網羅的に評価する必要がある。そのため、本ガイドラインでは①データの設計、②データの品質、③学習済みモデルの品質、④実装・運用の品質、の区分にそれぞれ 2 つずつの内部品質を設定し、それらのマネジメントを通じて外部品質を実現することを求める。これらの内部品質は全ての機械学習要素に共通であり、利用時品質・外部品質と異なり、読者が自ら設定する必要は無い（利用時品質の設定については 2.2.1、外部品質の設定については 2.2.2 を参照）。

「機械学習品質マネジメントガイドライン 第 1 版」では、それぞれの内部品質の確保するために実施すべき事項を要求事項としてとりまとめており、要求事項は、外部品質のレベルに応じた 3 つのレベルに分類されている。必要なレベルに該当する要求事項に従って機械学習要素を作り込むことを通して、外部品質を満たす機械学習要素を実装する（外部品質を満たしていることは、テストにより確認する）。内部品質のレベルの確認については 2.2.4 で、内部品質の要求事項の確認および実行については 2.2.5 で詳述する。

本ガイドラインでは、「機械学習品質マネジメントガイドライン 第 1 版」の要求事項をそのまま用いる。補足として、その分野横断的な要求事項をプラント保安分野において適用する際に踏まえる必要がある「観点」を「附録 内部品質確保のための『プラント保安分野での観点』チェックリスト」にまとめている。ここには、「AI プロダクト品質保証ガイドライン 2020.08 版」（1.2.1(2)に詳述）の 7 章「産業用プロセス」ドメインで挙げられている留意事項も反映している。また、プラント保安分野の典型的なユースケース特有の「観点」

も、上記の附録に記載している（ユースケースについては3章で詳述）。

以下、図 2-1 で整理されている「機械学習品質マネジメントガイドライン 第1版」の内部品質の8つの軸について、「データの設計」「データの品質」「学習済みモデルの品質」「実装・運用の品質」の分類ごとに解説する。なお、要求事項ここでは例のみを記載しており、全体は上記の附録にまとめている。

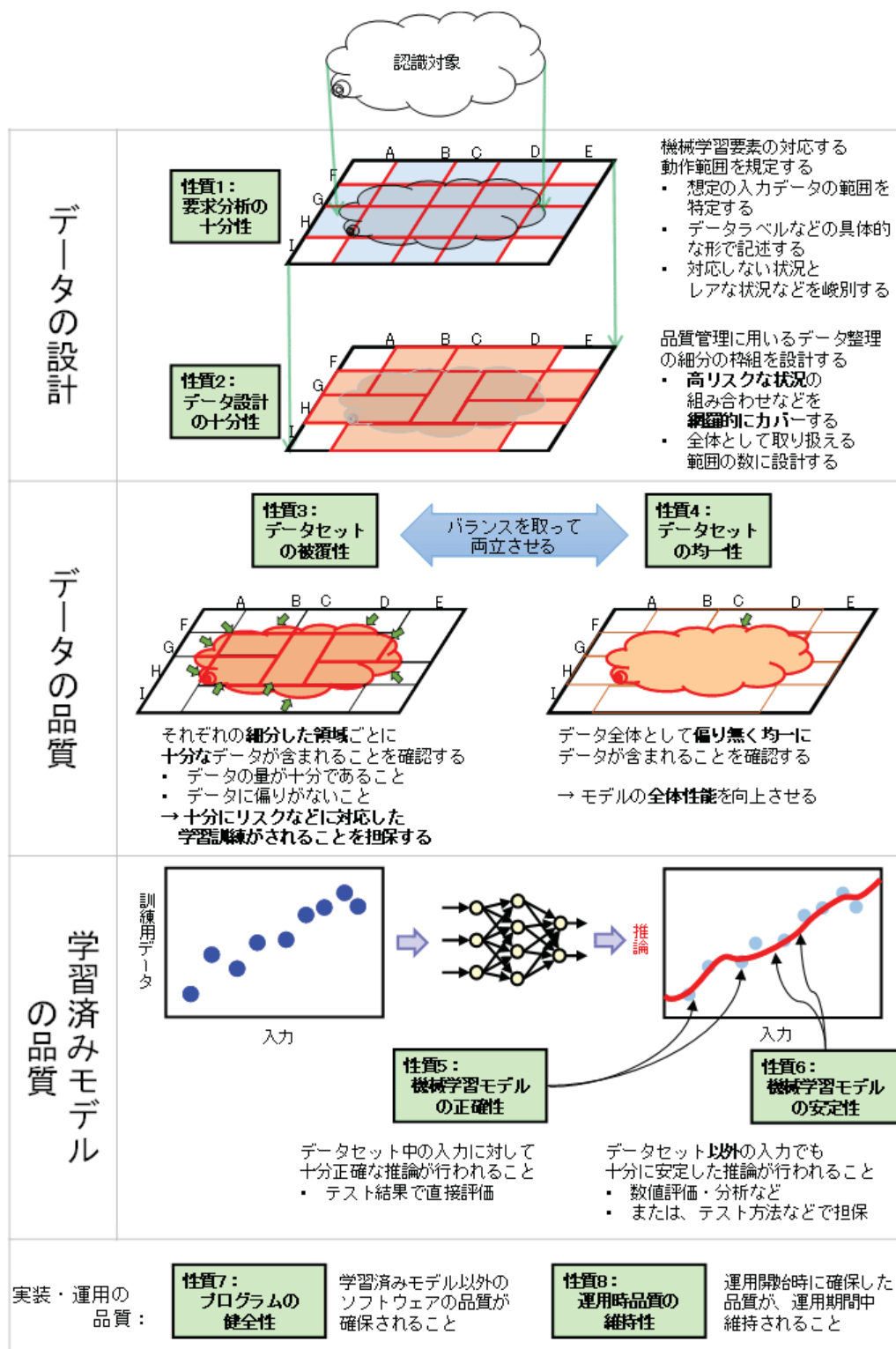


図 2-2 内部品質の 8 つの軸及びその関係性

【データの設計に関する軸】

(1) 要求分析の十分性

「機械学習品質マネジメントガイドライン 第1版」における定義

機械学習利用システムの実世界での利用状況に対応して機械学習要素に入力されると想定される運用時の実データの性質について分析が行われ、その分析結果が想定される全ての利用状況を被覆していること。

「機械学習品質マネジメントガイドライン 第1版」における要求事項の例

- ✓ 機械学習要素が対応すべき動作の対象・範囲を明らかにすること
- ✓ 機械学習要素の入力データの範囲を特定すること
- ✓ 機械学習要素が対応しない状況、発生頻度が低い状況などを判断すること
- ✓ 機械学習要素に起因する、機械学習利用システムの品質が低下するリスクを検討すること

(解説)

機械学習要素が対応する動作範囲を規定すること。想定する入力データの範囲を特定し、データラベルなどの具体的な形で記述し、対応しない状況やレアな状況等を峻別する。

(2) データ設計の十分性

「機械学習品質マネジメントガイドライン 第1版」における定義

要求分析の十分性を前提として、システムが対応すべき様々な状況に対して十分な訓練用データやテスト用データを収集し整理するためのデータ設計の十分な検討を実施すること。

「機械学習品質マネジメントガイドライン 第1版」における要求事項の例

- ✓ 機械学習利用システムの品質が低下するリスクを有するケースを網羅すること
- ✓ 取り扱える範囲のデータの属性及びデータ量とすること
- ✓ ケースが少ない場合は網羅して確認し、多い場合は、いくつかのケースを属性や組合せを網羅するように取り出し確認すること
- ✓ 高い品質が要求される場合は、ケースの抽出作業に数学的な「網羅性基準」を導入すること

(解説)

品質管理に用いるデータ整理の細分の枠組を設計すること。高リスクな状況の組み合わせなどを網羅的にカバーすることや、全体として取り扱える範囲の数に設計すること等が含まれる。

【データの品質に関する軸】

(3) データセットの被覆性

「機械学習品質マネジメントガイドライン 第1版」における定義
対応すべき状況の組み合わせの各々に対して、状況の抜け漏れがなく、十分な量の学習データが与えられていること。
「機械学習品質マネジメントガイドライン 第1版」における要求事項の例
✓ データセットの取り方などを工夫し、必要なデータを網羅的に確保すること ✓ 発生頻度の低いデータが得られない場合は、テスト等も通じて検証し個別に対応を検討すること ✓ 網羅性基準を導入している場合、「ケースに含まれていない」属性が偏りなく分布しているか検査すること

(解説)

細分した領域ごとに、十分なデータが含まれることを確認すること。データの量が十分あり、偏りが無いことを確認する。これにより、十分にリスクなどに対応した学習訓練がされることを担保する。

(4) データセットの均一性

「機械学習品質マネジメントガイドライン 第1版」における定義
データセット内の各状況や各ケースが、入力されるデータ全体におけるそれらの発生頻度に応じて抽出されていること。
「機械学習品質マネジメントガイドライン 第1版」における要求事項の例
✓ データセット全体を取得するプロセスに偏りが生じないよう配慮しつつ、個々の属性値の発生頻度などを適宜監視すること ✓ 「データセットの被覆性」とどのように両立させるのかの検討やデータ設計を行うこと

(解説)

データ全体として偏りなく均一にデータが含まれることを確認すること。これはモデルの全体性能の向上を意図している。

「データセットの被覆性」と「データセットの均一性」は、バランスをとって両立させることが必要である。

なお、「データセットの均一性」は、「リスク回避性」のレベルと「パフォーマンス」のレベルに応じてそれぞれ要求事項が設定されている点に留意が必要である。

【学習済みモデルの品質に関する軸】

(5) 機械学習モデルの正確性

「機械学習品質マネジメントガイドライン 第1版」における定義

学習データセット（訓練用データ、テスト用データ、バリデーション用データからなる）に含まれる具体的な入力データに対して、機械学習要素が期待通りの反応を示すこと。

「機械学習品質マネジメントガイドライン 第1版」における要求事項の例

- ✓ データセットの入力データに対して、機械学習要素が外部品質の期待に沿った出力であること
- ✓ データの量を変化させたり、交差検定の手法を用いたりするなど学習の達成度を評価すること
- ✓ 出力に対し一定の誤判断を許容する場合は、その判断基準を定めること

（解説）

データセットの入力に対して十分正確な推論が行われること。テストの結果により、訓練用データやテスト用データの精度が直接評価される。

(6) 機械学習モデルの安定性

「機械学習品質マネジメントガイドライン 第1版」における定義

学習データセットに含まれない入力データに対して、機械学習要素が学習データセット内のそれに近いデータに対する反応と十分に類似した反応を示すこと。

「機械学習品質マネジメントガイドライン 第1版」における要求事項の例

- ✓ データセットに含まれない入力データに対して、その入力データ近傍のデータセットを入力したときの出力と類似していること

（解説）

データセット以外の入力データに対しても十分に安定した推論が行われること。数値評価・分析やテスト方法などで担保する。特に安全性が要求される場合、運用中のデータに対して安定した性能が得られる必要があり、「機械学習モデルの安定性」の確保は重要である。

【実装・運用の品質に関する軸】

(7) プログラムの健全性

「機械学習品質マネジメントガイドライン 第1版」における定義

機械学習の訓練段階に用いる訓練用プログラムや、実行時に使われる予測・推論プログラムが、与えられたデータや訓練済み機械学習モデルなどに対してソフトウェアプログラムとして正しく動作すること。

要求事項の例

- ✓ 利用するソフトウェアは信頼できるソフトウェアを用いること
- ✓ 機械学習要素の開発環境と運用環境の相違とそれによる影響を予め検討すること

(解説)

機械学習モデル以外のソフトウェアの品質が確保されること。機械学習要素の開発に用いるソフトウェアなどが健全であることが求められる。また、機械学習要素以外の構成要素は、一般的なシステムに求められる品質や要求事項に則して開発することを前提とする。

(8) 運用時品質の維持性

「機械学習品質マネジメントガイドライン 第1版」における定義

運用開始時点で充足されていた内部品質が、運用期間中を通じて維持されること。

要求事項の例

- ✓ 機械学習要素の更新頻度の見積もりまたは更新の必要性の判断基準を事前に検討すること
- ✓ 更新時の品質検査の方法、特に更新の可否の判断基準（または意思決定の方法）について検討すること

(解説)

運用開始時に確保した品質が、運用期間中維持されること。運用時品質を維持するために予め検討すべき事項が要求事項となる。特に、プラントは製品の切り替えやメンテナンス等様々な要因で経時変化する。機械学習利用システムを運用することで様々な設備状態の変化や新たなデータが得られることにより、当初の内部品質が維持されず「リスク回避性」「パフォーマンス」が維持されないことにもつながりかねないため、「運用時品質の維持性」の確保は重要である。

2.2 信頼性評価の方法と要求事項

本節では、機械学習要素の信頼性評価の適用方法や要求事項について、3段階の品質（利用時品質、外部品質、内部品質）に従って詳述する。信頼性評価を行う際には、本節の手順に則って品質項目の具体化やレベルの設定を行う。適用方法の説明に用いる3段階の品質の関係図を以下に示す。

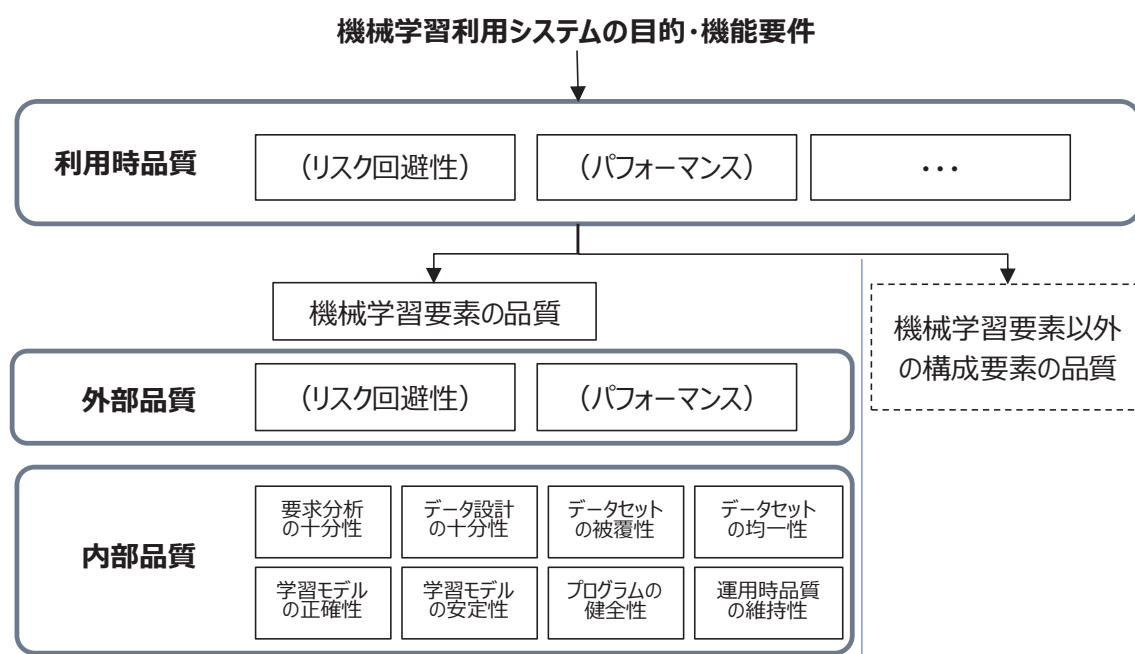


図 2-3 信頼性評価における3段階の品質の関係

2.2.1 利用時品質の設定

機械学習利用システムの機能要件（実現したいこと）を踏まえて、利用者の視点で利用時品質を設定する。

機械学習利用システムの機能要件から、利用時品質を設定するイメージを図 2-4 に示す。³⁶例えば「異常予兆検知・診断」を行う機械学習利用システムの場合、システムの出力をプラントの運転を担う運転員が確認し、プラントに対する必要な操作を行う。よって、利用時品質は、運転員が求める品質であり、「様々なプラントの条件下で、将来の異常の発生を正しく検知する」こと（「リスク回避性」の軸に該当）、「発報内容の確認に運転員や点検要員の時間を割かなくてよい適度な発報頻度とする」こと（「パフォーマンス」の軸に該当）を求めることが考えられる。

³⁶ 3 節では、ユースケース別の利用時品質の項目例を記載している。

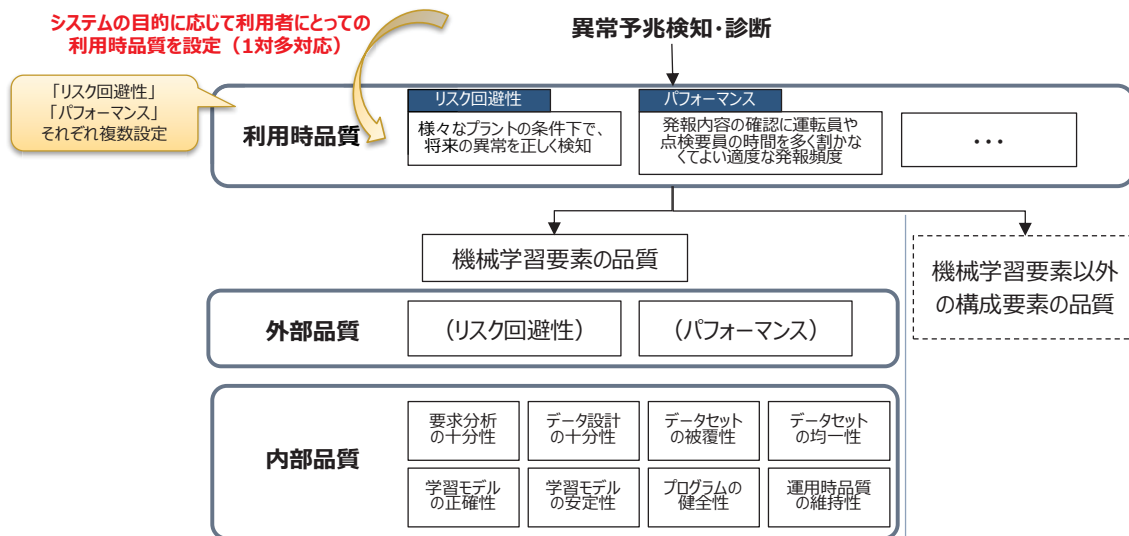


図 2-4 利用時品質の設定イメージ（異常予兆検知・診断システムの例）

2.2.2 外部品質の設定

前節で設定した利用時品質に対応した機械学習要素の外部品質を設定する。外部品質は機械学習要素の出力に対して求められる品質であり、利用時品質に通常 1 対 1 対応で設定される。2.2.2 の時点では外部品質の設定段階では機械学習特有の数値目標（例：〇%以上の精度）を定める必要はない。「リスク回避性」の外部品質に関しては、外部品質のレベル設定(2.2.3)の過程で、安全関連系や外部安全機構を確認することを通して機械学習要素に要求される数値目標（危険につながる誤判断の発生率等）を定めることができる場合がある。また、最終的には、機械学習要素を作り込む段階（2.2.5）で、PoC の成果やデータの取得状況・学習状況等に応じて具体的な機械学習特有の数値目標（正解率(Accuracy)、F 値など）を設定できる。

機械学習利用システムの利用時品質から、機械学習要素の外部品質を設定するイメージを図 2-5 に示す。例えば「異常予兆検知・診断」を行う機械学習利用システムの場合、「様々なプラントの条件下で、将来の異常の発生を正しく検知する」という利用時品質に対しては、「『異常予兆あり』の場合に、『正常』であると判定する誤判定率を限りなく小さくする」という外部品質が対応する（「リスク回避性」の軸に該当）。同様に、「発報内容の確認に運転員や点検要員の時間を多く割かなくてよい適度な発報頻度とする」という利用時品質に対しては、「誤検知の数をできるだけ少なくする」という外部品質が対応する（「パフォーマンス」の軸に該当）。

利用時品質は、「機械学習要素」と「機械学習要素以外の構成要素」それぞれの外部品質から成り立つものであるが、本ガイドラインでは機械学習要素の外部品質のみを対象としている。

なお、プラント保安分野の場合、機械学習利用システムと独立の安全関連系が確立している場合は、機械学習利用システムについて「リスク回避性」の考慮が不要になる場合もある。例えば、設備の劣化傾向を検知する機械学習利用システムについて、設備の部品のブレ（変位）が一定以上になっていることを検知するセンサなどが別途存在し、安全面はその独立した安全関連系によって担保される場合、機械学習利用システムの目的は、劣化傾向を早期に

把握して効率的なメンテナンス計画を作成することに絞られる。このため、「リスク回避性」の考慮が不要になる（「3.3.3 設備劣化診断」のユースケース参照）。

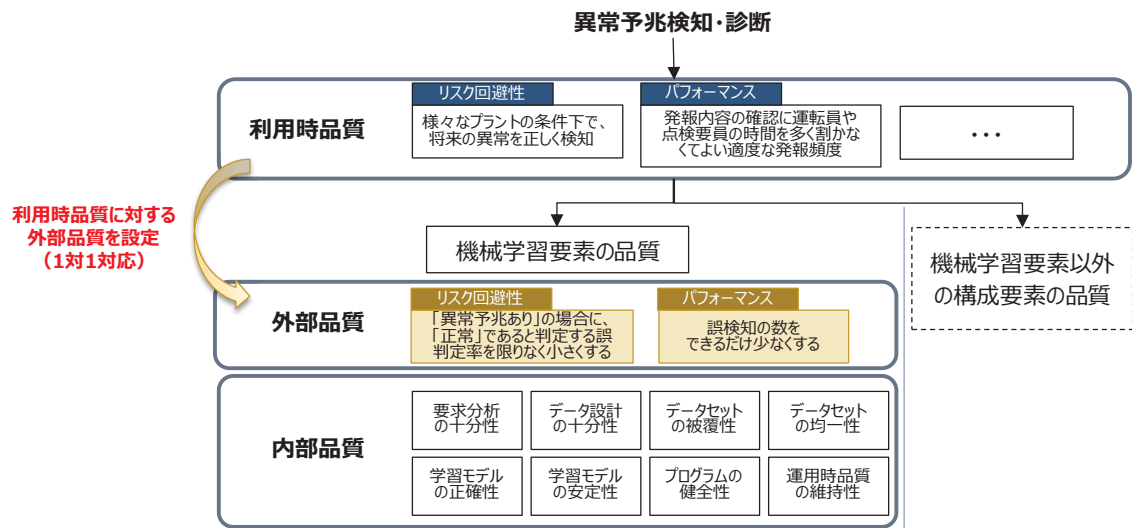


図 2-5 外部品質の設定イメージ（異常予兆検知・診断システムの例）

2.2.3 外部品質のレベル設定³⁷

設定した外部品質について、機械学習要素への要求に応じたレベルを設定する。「機械学習品質マネジメントガイドライン 第1版」の定義を用いて、「リスク回避性」は AISL (AI Safety Level)、「パフォーマンス」は AIPL (AI Performance Level) という指標でレベルを設定する。

(1) リスク回避性

「リスク回避性」軸の外部品質に対し、機械学習要素の誤判断による悪影響を回避することへの要求の強さに応じてレベルを設定する。「リスク回避性」に関する外部品質のレベル設定の方法を以下に詳述する。

1) レベル設定の手順

リスク回避性のレベル(AISL)の設定のフローを図 2-6 に示す。

³⁷ 外部品質のレベル設定は、IEC61511-1(JIS C 0511-1)の「プロセス潜在危険及びリスク評価 (H&RA)」 「安全機能の防護層への割当て」に対応するステップである。ただし、本ガイドラインの適用が、IEC61511-1(JIS C 0511-1)への適合を意味するものではない。

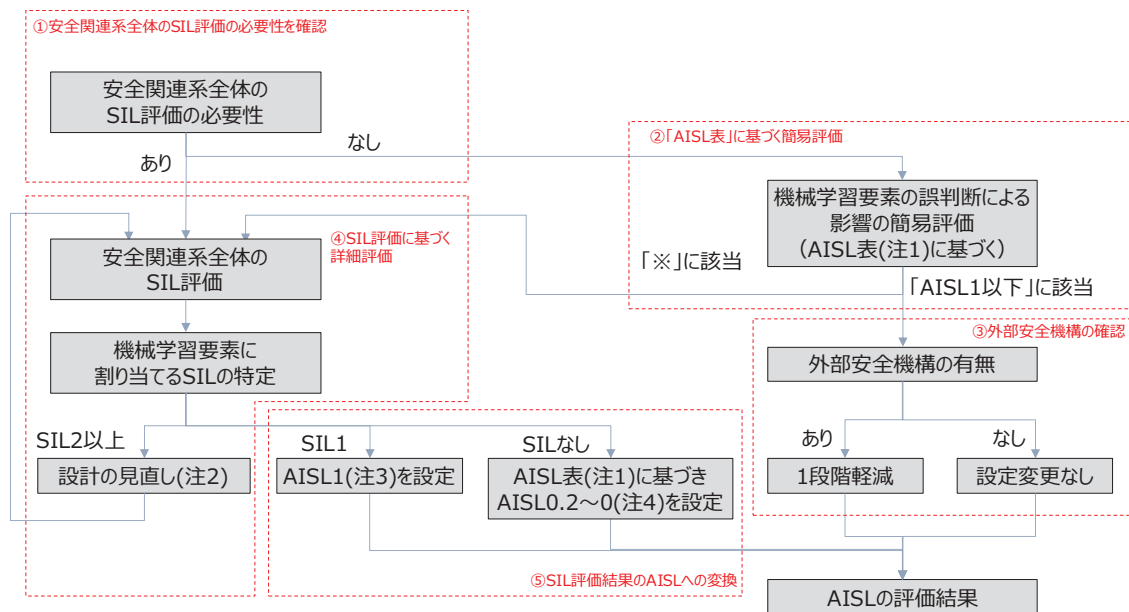


図 2-6 リスク回避性のレベル(AISL)の設定フロー

注 1) 「AISL 表」とは、表 2-1 を指す。

注 2) 本ガイドラインでは、外部安全機構により、機械学習要素への SIL の割当を SIL1 以下とする。

注 3) SIL 評価に基づく SIL1 以上の安全度を担保するには、IEC 61508(JIS C 0508)への適合が必要。

「AISL1 を満足する機械学習要素を SIL1 として使用できる」ことは意味しない。

注 4) 「SIL なし」に相当する AISL は「0.2~0」である。機械学習要素に割り当てる SIL が「SIL なし」であり、かつ AISL 表に基づく AISL 評価が 1 以上となる場合は、より厳密な評価である「SIL 評価」を優先する。すなわち、AISL 表の評価が「※」「AISL1」であっても、SIL 評価による「SIL なし」を優先し、AISL は「0.2」を設定する。AISL 表の評価が 0.2~0 となった場合は、AISL 表の評価結果を用いる。

①安全関連系全体の SIL 評価の必要性を確認

まず、機械学習利用システムを実装する対象について、安全関連系全体の SIL 評価の必要性を確認する。SIL 評価とは、機能安全規格³⁸において定められている、安全関連系の安全度水準(SIL)の要求を定める手法であり、SIL 評価に基づいて設計された安全関連系を実装することで、安全確保の手段が適切であることを示すことができる。機械学習利用システムを実装する設備が、機能安全規格の適用対象と想定される場合は、SIL 評価を要すると判断する。これ以外の場合は SIL 評価を要しないと判断する。

また、機能安全規格の適用対象となる設備であっても、機能安全規格³⁸に従ったシステム開発プロセスの手法により信頼性が確認された他の安全関連系（機械学習利用システムと独立したシステム）によってプラントの安全は確保されており、機械学習要素が誤判断をしても安全に何も影響しないことが確認されている場合は、新規の SIL 評価は要しないと判断する。

②「AISL 表」に基づく簡易評価

SIL 評価を要しないと判断された場合は、表 2-1(AISL 表)に基づいて、機械学習要素への

³⁸ IEC61508(JIS C 0508)、IEC61511(JIS C 0511)等。

リスク回避の要求の強さを簡易評価する(表の取り扱いについては後述)。ここで、表 2-1 内の「※」に該当する場合は、機械学習要素へのリスク回避の要求が過大になっている(現時点の機械学習技術と信頼性のマネジメント技術では達成できないほど高いリスク回避レベルを要求している)恐れがあり、簡易な手法による評価が妥当でない可能性があるため、改めて SIL 評価が必要と判断する。

③外部安全機構の確認

AISL1 以下に該当する場合で、外部安全機構³⁹が存在しない場合は、該当する値をそのまま AISL の評価値とする。

外部安全機構が存在する場合は、該当する値から 1 段階(例：1→0.2、0.2→0.1)軽減した値を AISL の評価値⁴⁰とする。ただし、外部安全機構は、機械学習要素の望ましくない出力及び判断に対して監視・補正(それらの制限や上書き修正)を行い、かつ、SIL 評価に基づいて設計された常時稼働するものでなければならない。⁴¹本ガイドラインでは、外部安全機構があれば、機械学習要素の出力及び判断に対して、適切に監視・補正(出力の制限や上書き修正)が行われるため、AISL を軽減させることができるとしている。

④SIL 評価に基づく詳細評価

SIL 評価が必要と判断された場合は、機能安全規格に従って安全関連系の SIL 評価を行い、機械学習要素に割り当てる SIL を特定する。本ガイドラインにおいては、2020 年 11 月時点の機能安全に係る国際的な議論の動向を踏まえ、SIL2 以上の安全機能を機械学習要素に割り当てることは安全性へのリスクが大きいため、不可とする⁴²。よって、機械学習要素に割り当てられた SIL が 2 以上となった場合は、設計を見直し、機械学習要素の SIL が 1 以下となるようにする。なお、機械学習要素の外部品質の数値目標(危険につながる誤判断の発生率等)を設定する際に、機械学習要素に割り当てる SIL に対応した目標機能失敗尺度(例：SIL1 の低頻度作動要求モードのとき、作動要求あたりの機能失敗平均確率が 10^{-2} 以上 10^{-1} 未満)を参考に検討することも考えられる⁴³。

³⁹ 安全性を向上させることを目的として機械学習要素と並列・直列に処理されるソフトウェアまたはハードウェアで、機械学習要素の望ましくない出力に対して監視・補正(出力の制限や上書き修正)を行い、かつ IEC 61508(JIS C 0508)、IEC 61511(JIS C 0511)などの機能安全規格に従う既存のシステム開発プロセスの手法で十分に安全を確保できると評価できるもの。

⁴⁰ AISL の評価値は、1/0.2/0.1/0 としている。AISL1 は SIL1、0.2~0 は SIL なしと対応する。SIL 評価は機能安全規格(IEC61508/JIS C 0508)に基づき 4/3/2/1 の 4 段階に分類されており、分類に応じた安全機能の要求が規定されている。「SIL なし」の場合は、機能安全規格に基づく特別な要求事項の規定はないが、通常の品質管理が求められる。しかし、機械学習は従来のシステムのように通常の品質管理の手法が確立していないため、一定の指針が必要である。そのため、「SIL なし」としているレベルをさらに分割し、大小関係を保つために小数を用いて 0.2、0.1、0 というレベル表記を採用し、一定のリスク回避性を求める設定としている。(「機械学習品質マネジメントガイドライン 第 1 版」における定義を継承)

⁴¹ 外部安全機構がソフトウェアの場合は、機械学習要素との独立性について特に留意が必要である。

⁴² 今後、ガイドラインを改訂する際に、最新の機能安全に係る議論の動向を踏まえて更新する。

⁴³ IEC 61508-1(JIS C 0508-1)において定められている、SIL の水準に対応した機能失敗平均確率(例：SIL1 の低頻度作動要求モードのとき、作動要求あたり 10^{-2} 以上 10^{-1} 未満、等)が参考となる。詳細は、IEC 61508-1(JIS C 0508-1)を参照。ただし、ソフトウェアに適用する場合は、IEC 61508-3(JIS C 0508-3)におい

⑤SIL 評価結果の AISL への変換

機械学習要素に割り当てる SIL が 1 の場合は、AISL は 1 とする。

SIL なしの場合は、表 2-1(AISL 表)に基づいて、AISL を設定する。ただし、評価の結果が「※」又は「AISL1」となった場合は、より厳密な評価である「SIL 評価」の結果を優先して適用することとし、「SIL なし」に対応する AISL の 0.2、0.1、0 のうち、最も安全性に配慮した AISL0.2 を設定する。

2) 「AISL 表」

表 2-1(AISL 表)は、機械学習要素の誤判断に起因する人的または経済的な影響の強度を縦軸、機械学習要素の誤判断を人間が回避する可能性の程度を横軸にとっている。これは、「機械学習品質マネジメントガイドライン 第 1 版」記載の AISL の評価表を基にして、縦軸・横軸をプラント保安分野の実態を考慮して改変したものである。

て、機能失敗平均確率ではなく他の技法が SIL ごとに要求されている。これは、ソフトウェアに対して機能失敗平均確率の要求を設定することが困難とされているためである。このため、ここで設定した数値目標は、内部品質の作り込みの段階でも、PoC 等を通して調整されることが想定される。

表 2-1 「リスク回避性」の簡易評価の基準（AISL 表）

	人間による回避可能性の分類 ⁴⁴		
	①人間による代替システムが用意されておらず、機械学習要素の判断結果がそのまま運転・保全に反映される	②機械学習要素の判断結果が運転・保全にそのまま反映されることはなく、人間による確認・代替システム適用を介した結果が反映される	③機械学習要素は補助的情報のみを提示し、人間が判断して運転・保全に反映する
I：死亡・障碍の残る傷害・多数の重傷・極めて多数の負傷・直接被害額1億円以上	※（注1）	※（注1）	※（注1）
II：重軽傷・多数の負傷・直接被害額1千万円以上	※（注1）	AISL 1(注3)	AISL 0.2(注3)
III：微小な怪我(注2)・直接被害額1千万円未満	AISL 1(注3)	AISL 0.2(注3)	AISL 0.1
III'：(想定される被害者により容易に回避できる場合)	AISL 0.2(注3)	AISL 0.2(注3)	AISL 0.1
IV：傷害の想定なし・直接被害額の想定は軽微	AISL 0	AISL 0	AISL 0

注1) 本ガイドラインでは、※箇所に該当する場合、安全関連系全体の SIL 評価を必須とし、機械学習要素に割り当てる SIL が SIL1 以下となるよう設計する。

注2) 「微小な怪我」とは、いわゆる「赤チン災害」以下の強度にあたり、医師の診察を要する不休災害などは強度基準Ⅱに該当する。

注3) ※箇所に該当しない場合で、かつ SIL 評価に基づいて設計した常時稼働する外部安全機構が存在する場合、記載されている AISL から1段階(例：1→0.2、0.2→0.1)軽減できる。

注4) 本表では、安全性を重視した簡易的な評価とするため、リスクグラフ法等に基づく SIL の設定で考慮される発生確率のランクを設けず、一律発生確率が高いものとして取り扱っている（＝SIL 設定の際に考慮する発生確率が最も高いものとして、その SIL 値に相当する AISL を割り振っている）。

a. 「AISL 表」の縦軸

「AISL 表」の縦軸は、機械学習要素が誤判断した場合に発生する人的または経済的な影響の強度に応じて AISL が変動することを示している。したがって、強度が大きい表の上の方が求められる AISL が高く、下の方が低く設定されている。この軸に則って AISL を設定する。

表 2-2 は、本ガイドラインで定めた人的または経済的な影響と、「機械学習品質マネジメントガイドライン第1版」やプラント保安分野の既存の評価基準・事故分類との関係を整理したものである。表 2-1（AISL 表）に基づいて縦軸の強度基準を検討する際の参考とされたい。

⁴⁴ 「機械学習品質マネジメントガイドライン 第1版」における回避可能性の分類を、プラント保安分野として解釈したもの。

表 2-2 本ガイドラインにおける強度基準と既存の基準・分類との関係

「想定される影響」の既存評価基準			「発生した事故」の既存評価基準		
機械学習品質マ ネジメントガイドラ イン 第1版	【高圧ガス】 リスクアセスメント ガイドライン(Ver. 2) リスクアセスマ ント実施事例	【労働安全】 危険性又は有害性等 の調査等に関する指針 別添4 負傷または疾 病の重篤度	【高圧ガス】 高圧ガス・石油コンビナート事 故対応要領 事故の分類	【消防】 危険物施設における火 災・流出事故に係る深 刻度評価指標 人的被 害指標	【石化協】 事故評価基準 (CCPS評価法)
Ⅰ ・死亡 ・障碍の残る傷害 ・多数の重傷 ・極めて多数の負傷(重傷・ 軽傷) ・直接被害額 1 億円以上	複数人の同時死 亡	Ⅰ：死亡 <			

b. 「AISL 表」の横軸

「AISL 表」の横軸は、機械学習要素の誤判断を人間が回避する可能性の程度に応じて AISL が変動することを示している。①は、機械学習要素の判断がそのまま機械学習利用システム全体の最終判断となる場合が該当する。②は、機械学習要素は何らかの判断を行うが、その判断は必ず人間が確認し、最終判断となる場合が該当する。③は、機械学習要素自体は判断を行わず（判断の材料となる情報のみを出力し）、機械学習要素の出力を必ず人間が見て判断を行う場合に該当する。したがって、表の左の方が求められる AISL が高く、右の方が低く設定されている。この軸に則って AISL を設定する。

横軸の検討にあたっては、図 2-7 に基づき、機械学習要素の出力及び判断に対する人間の関与の度合いを確認する。

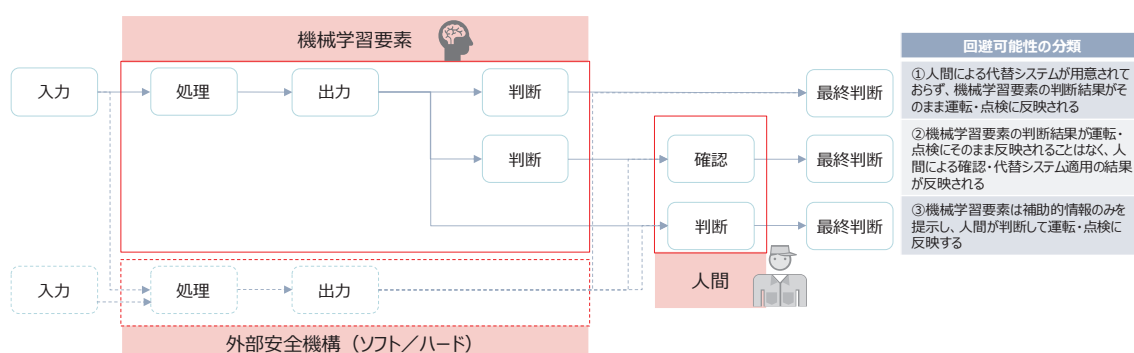


図 2-7 機械学習要素の出力・判断に対する人間の関与

(2) パフォーマンス

「パフォーマンス」に関する外部品質のレベル設定及び評価基準を表 2-3 に示す。産業技術総合研究所が策定した「機械学習品質マネジメントガイドライン 第 1 版」に示されている内容を継承し、AIPL という基準で評価する。AIPL は、基本的には、「パフォーマンス」に関する外部品質が、システムにおいてどの程度強い要件となっているかで決定する。

最も大きいレベルである AIPL 2 は、機械学習要素が一定の性能指標を満たすことが、機械学習利用システムの運用上必須または強い前提である場合に相当する。また、AIPL 1 は、一定の性能要件が機械学習利用システムの目的として特定されているが、AIPL 2 に該当しない場合（性能要件の達成がベストエフォートでよく必達ではない場合）に相当する。また、AIPL 0 は、性能指標が特定されず、性能指標そのものを発見することが開発の目的となる場合などに相当する。この AIPL2・1・0 の評価基準は、「機械学習品質マネジメントガイドライン第 1 版」において示されている基準と同一である。

表 2-3 「パフォーマンス」のレベル設定・評価基準（AIPL）

パフォーマンスレベル	説明
AIPL 2 (mandatory requirements)	・機械学習要素が一定の性能指標（正答率・適合率・再現率など）を満たすことが、機械学習利用システムの運用上の必須または強い前提である場合。 ・受発注等の契約において、前記の一定の性能指標の充足が受入要件として明確に記載される場合。
AIPL 1 (best-effort requirements)	・一定の性能要件が機械学習利用システムの目的として特定されているが、AIPL 2 に該当しない場合。 - 特に、リリースまでの日程スケジュールが重視される場合、または品質をモニタリングしながら試験運用を行い、漸次性能向上を行う運用が許される場合。
AIPL 0	・性能指標が開発時点で特定されず、性能指標そのものを発見することが開発の目的となる場合など。 ・所謂 PoC の段階で終了する開発を行う場合。

なお、「機械学習品質マネジメントガイドライン 第1版」における「パフォーマンス」の評価基準は、「性能（正答率・適合率・再現率など）が一定の水準を満たすことに対する要求の厳しさ」によってレベルを設定しており、「性能の水準そのもの」のレベルは示していない。本来は「性能の水準そのものが高い場合に、より厳しいレベルが要求される」ところであるが、「機械学習品質マネジメントガイドライン第1版」の基準では、性能に求められる水準の高低によらずレベルは同じ値で評価される形となっている⁴⁵。

本ガイドラインでは、プラント保安分野への適用にあたり、実用的な「パフォーマンス」の評価を行う基準とするため、「ベストエフォートで運用すればよいが、性能の水準の要求が高い」場合があれば「AIPL2」を適用すべきこととする。⁴⁶

ここまで説明した AIPL の考え方を図 2-8 に示す。

⁴⁵ 「機械学習品質マネジメントガイドライン 第1版」は、達成すべき具体的な目標値そのものは応用ごとに異なるため、性能の水準そのものを AIPL の指標に含めていない。本ガイドラインはプラント保安分野に特化しており、具体的応用を想定できることから、性能の水準そのものを AIPL の指標に含めることとした。

⁴⁶ 本ガイドラインのユースケース「配管の肉厚予測」において、性能の要求水準が高い場合の AIPL2 の適用について例示している（3.3.1 2）b.を参照）。

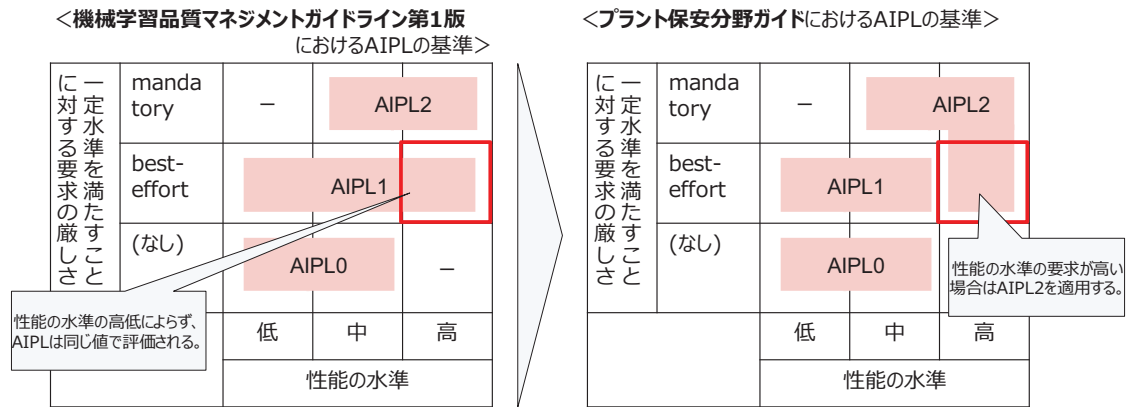


図 2-8 「機械学習品質マネジメントガイドライン 第1版」と本ガイドラインのAIPLの考え方

2.2.4 内部品質のレベルの確認

前項でレベルを設定した外部品質は、内部品質を作り込むことにより達成する。前項で設定した AISL/AIPL に応じて、内部品質の要求水準（要求レベル）を設定する⁴⁷。外部品質の AISL/AIPL と内部品質の要求レベルとの対応関係を図 2-9 に示す。例えば、AISL 0.2 もしくは AIPL 2 の場合、内部品質の要求は「レベル2」となり、内部品質の8つの軸それぞれに設定された「レベル2」の要求事項を適用することとなる。複数の AISL/AIPL について、それぞれ内部品質のレベルを設定し、最も高い内部品質のレベルを用いて要求事項を適用する。

外部品質のレベル				
AISL 0.1	AIPL 1			
AISL 0.2	AIPL 2			
AISL 1	-			
内部品質の軸		レベル1	レベル2	レベル3
要求分析の十分性		✓ 主要な品質低下リスクが発生する原因について検討を行い記録する。 ✓ ...	✓ システム全体での利用時品質低下リスクとその影響について、工学的に一定の網羅性を持つ分析を行い、文書として記録する。 ✓ ...	✓ レベル2に加えて、以下の活動を行う。 ✓ ...
データ設計の十分性		✓ ... ✓ ...	✓ ... ✓ ...	✓ ... ✓ ...
...		✓ ... ✓ ...	✓ ... ✓ ...	✓ ... ✓ ...

図 2-9 外部品質の AISL/AIPL と内部品質の要求レベルの対応

2.2.5 内部品質の要求事項の確認および実行

前項で確認した内部品質のレベルに基づく要求事項に従って、機械学習要素の開発を行

⁴⁷ 「データセットの均一性」のみ、AISL と AIPL にそれぞれ独立に対応したレベル（AISL0.1→LvS1、AISL0.2,1→LvS2、AIPL1→LvE1、AIPL2→LvE2）が設定されている。附録 内部品質確保のための「プラント保安分野での観点」チェックリストの「データセットの均一性」を参照。

う。その際には①「要求事項」、②「プラント保安分野での観点」、③「ユースケース固有の観点」を確認する必要がある。⁴⁸

内部品質の8つの軸及びレベル別の要求事項は、「機械学習品質マネジメントガイドライン 第1版」の要求事項を用いる(①)。

更に本ガイドラインでは、要求事項を実際のプラント保安分野における機械学習要素の開発等に円滑に適用できるよう、プラント保安分野における開発の実例に基づき、プラント保安分野特有の留意点を整理した(②)。また、本次章で説明する5件のユースケースを設定し、ユースケースに固有の留意点も整理している(③)。

①②③は「附録 内部品質確保のための『プラント保安分野での観点』チェックリスト」にまとめており、機械学習要素の構築・運用の際には附録を参照し、「①要求事項」と併せて必要に応じて「②プラント保安分野での観点」「③ユースケース固有の観点」を確認しながら、要求事項を実現する。

これにより、「機械学習要素の『内部品質』の向上を通じてその『外部品質』を必要となるレベルで達成し、最終的なシステムの『利用時品質』を実現する」⁴⁹という階層的な品質確保を実現する。

なお、外部品質の設定段階では機械学習特有の数値目標(例:○%以上の精度)を定めていないが、「リスク回避性」の外部品質に関しては、外部品質のレベル設定(2.2.3)の過程で機械学習要素に要求される数値目標(危険につながる誤判断の発生率等)を定めることができる場合があるほか、本項において機械学習要素を実際に作り込む中で、ユーザー企業・ベンダー企業双方の合意により、PoCの成果やデータの取得状況・学習状況等に応じて、具体的な機械学習特有の数値目標(正解率(Accuracy)、F値など)を設定することができる。

⁴⁸ ①②③は、本ガイドラインの「附録 内部品質確保のための『プラント保安分野での観点』チェックリスト」にまとめている。

⁴⁹ 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1版」

コラム: AIに「精度 100%」は必要？

機械学習は、データから規則性や判断基準を学習し、これに基づいて予測や判断を行うものである。どんなにデータを集めてもデータは現実のサンプルでしかないこと、学習した規則性や判断基準には決まったルールがないことなどから、100%の精度を達成することは原理的に困難である。

一方で、「安全」はプラントにとって至上命題であり、理念目標として「100%安全」を追求することは、プラントオーナーにとって基本的なスタンスである。そのため、プラントオーナーとしては、機械学習要素に対しても、その振る舞いに対して「100%の精度」を求めがちになり、開発目標の実現が困難となってプロジェクトが進まなくなってしまう場合がある。

本ガイドラインでは、機械学習要素単体で高いレベルの安全性を確保することは困難であるとの前提に立ち、安全性を既存のシステムとの組み合わせで確保することで、機械学習要素に過度な安全機能を任せないようにする、という考え方をとっている。

例えば、数ヶ月単位の中長期的な設備の劣化傾向を把握する機械学習利用システム（ユースケース「設備劣化診断(3.3.3)」）を導入する場合、このシステムの目的はあくまで中長期的なメンテナンスの最適化であり、「設備の故障を高精度に検知して安全を守ること」ではない。安全の機能は既存の安全関連系で担保されており、機械学習利用システムが判断を誤っても、機械学習要素導入以前と比べて安全が損なわれることはない。この場合、メンテナンスの最適化に資する精度さえあれば、システムを問題なく運用しメンテナンスを効率化することができる。

また、生産性を向上するための最適な運転パラメータを提示する機械学習利用システム（ユースケース「運転最適化(3.3.5)」）を導入する場合、機械学習要素の出力を監視して、異常な出力に対して運転パラメータの提示範囲を限定する「外部安全機構」や、インターロック等の既存のシステムと組み合わせて運用することにより、機械学習要素に対して極めて高いレベルの安全性の要求をする必要はなくなる。

加えて、数日～数十分後の短期的な異常の予兆を検知する機械学習利用システム（ユースケース「異常予兆検知・診断(3.3.4)」）を導入する場合、異常予兆のアラートから自動的にプラントを操作するのではなく、運転員が必要な確認を行ってから停止操作等を行う運用にすることで、機械学習要素に求める精度の水準を抑えることができる。

以上のように、機械学習要素単体ではなく、既存の安全関連系や「外部安全機構」、運転員・保全員を含めて総合的に安全を確保することが肝要である。これにより、機械学習の有用性を適切に活かし、保安力と効率性の双方を高めることができる。

次章からは、プラント保安分野における機械学習のユースケースをとりあげ、機械学習要素だけでなく総合的に安全を確保するための検討例を示す。これらを参考にして、自社の事例の安全確保策を考案し、機械学習要素の適切な精度目標を設定することが期待される。

3. プラント保安分野における機械学習のユースケース

3.1 本ガイドラインにおけるユースケースの位置づけ

前節で示した信頼性評価の手順は、機械学習の活用方法を問わず、プラント保安分野の全ての機械学習利用システムに適用できる。しかし、具体的な利用時品質や外部品質の設定や、内部品質の要求事項の実行を円滑に遂行するには、本ガイドラインへの一定程度の習熟を要するため、特に初めてガイドラインを活用する場合は、検討が煩雑になる懸念がある。

そこで、プラント保安分野の典型的な機械学習の活用事例ごとに本ガイドラインを活用する「ユースケース」を示して、利用時品質・外部品質の項目と内部品質の確保策を例示し、ガイドライン活用を検討する際に参考となる情報を提供する。

3.2 対象とするユースケース

本ガイドラインでは、保全（メンテナンス）に用いる機械学習利用システムとして「配管の肉厚予測」「配管の劣化診断」「設備劣化診断」の3件、運転（オペレーション）に用いる機械学習利用システムとして「異常予兆検知・診断」「運転最適化」の2件を取り上げる。各ユースケースの概要を表 3-1 に示す。

表 3-1 ユースケースの概要

ユースケース	導入目的	機能の概要
保全（メンテナンス）に用いる機械学習利用システム		
(1) 配管の肉厚予測	交換サイクルの適正化	配管の流量・内容物から配管の肉厚を予測する
(2) 配管の画像診断	目視点検の負荷軽減	配管の画像から目視点検の要否を判定する（スクリーニング）
(3) 設備劣化診断	交換サイクルの適正化	個別設備の部品の劣化を予測する
運転（オペレーション）に用いる機械学習利用システム		
(4) 異常予兆検知・診断	事故による運転停止の回避	プラントの異常予兆を検知する
(5) 運転最適化	生産効率・品質の維持・向上	目的に応じた最適な操作パラメータを提示する

図 3-1 は、プラントにおける事故等の進展と、本ガイドラインで設定したユースケースとの関係を示している。プラントの保安を事故等の進展（上部の青色のレーン）に対する防護機能ととらえた場合、防護機能を発揮する局面での活動としては、保全活動、運転操作、不適切操作対応、異常時対応、応急対応が挙げられる（中部の緑色のレーン）。それぞれの局面で機械学習の活用機会があると考えられるが、本ガイドラインの執筆時点では、保全活動と運転操作への活用を目的とした事例が多いため、ここに 5 つのユースケースを設定している。⁵⁰

⁵⁰ 今後、ガイドラインの改訂にあたり、プラント保安分野における機械学習利用システムの活用の進展を踏まえ、ユースケースの追加を検討する。

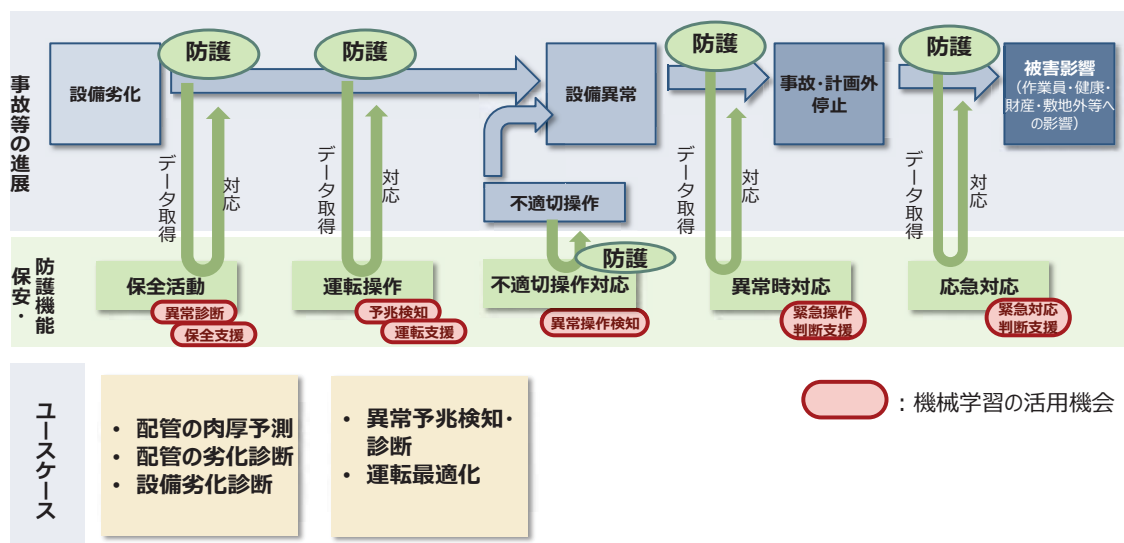


図 3-1 プラントにおける事故等の進展とユースケースの関係

3.3 ユースケースに基づく信頼性評価の具体的適用

本項では、各ユースケースについて、以下の構成で具体的な適用のイメージを示す。

まず「ユースケースの前提」として、例示する機械学習利用システムの機能や構成などの前提を定める。

次に「利用時品質・外部品質の項目例」として、当該ユースケースで検討する利用時品質及び外部品質の具体的な項目を、リスク回避性とパフォーマンスの軸に分けて例示する。外部品質の例示の中で「～（正しい判断の割合など）を『一定』以上にする」「～（誤検知の割合など）を『限りなく小さく』する」のように、『一定』『限りなく小さく』といった具体的な数字を想定した記載がある場合は、外部品質を設定する時点（2.2.2）では数値を決める必要はない。「リスク回避性」の外部品質に関しては、外部品質のレベル設定（2.2.3）の過程で、安全関連系や外部安全機構を確認することを通して機械学習要素に要求される数値目標（危険につながる誤判断の発生率等）を定めることができる場合がある。また、最終的には、機械学習要素を作り込む段階（2.2.5）で、PoC の成果やデータの取得状況・学習状況等に応じて、ユーザー企業・ベンダー企業双方の合意により、機械学習特有の具体的な数値目標（正解率(Accuracy)、F 値など）を設定する。

最後に「内部品質確保のための『ユースケース固有の観点』」として、内部品質の要求事項を満たす機械学習要素を開発するための当該ユースケース特有の留意点を提示する。

なお、ユースケースによらずにプラント保安分野で共通的に参照すべき留意点は、「附録 内部品質確保のための『プラント保安分野での観点』チェックリスト」にまとめている。本チェックリストは、内部品質を追求する際に活用できるよう、本ガイドラインが示している事項（要求事項とユースケースごとの観点）を一元的に示している。

各ユースケースの冒頭部には、プラント保安分野の事業者における、当該ユースケースに関連する実事例を示しており、ユースケースの理解のための参考とされたい。

なお、本ガイドラインでは、外部品質の AISL 及び AIPL の設定、ならびに、内部品質の要求レベルについては、ユースケースに基づいて示すのは外部品質の項目までであり、具体

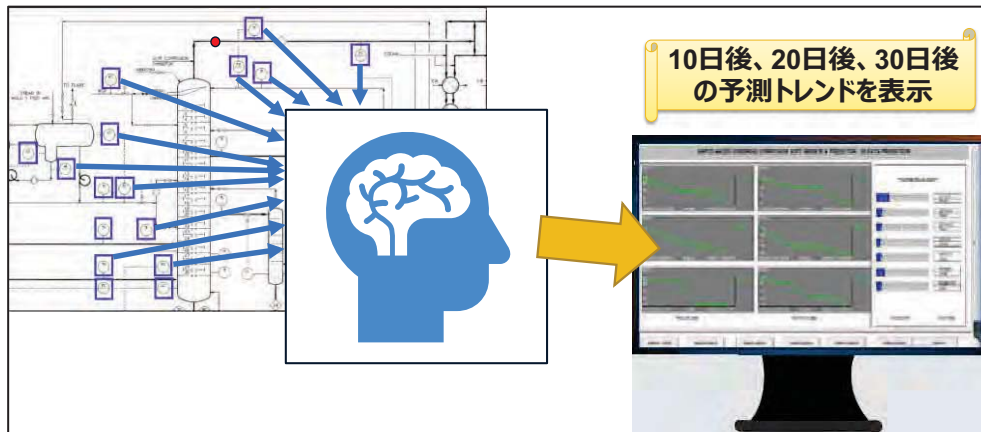
的な要求レベルの設定は行っていない。AISL/AIPL の設定・内部品質の要求レベルの設定は、実際の機械学習利用システム的前提や条件に基づいて、読者が自ら実施する必要がある。

また、本ガイドラインのユースケースは機械学習のモデル等に関して多くの前提を置いた上で、利用時品質や外部品質の具体化や内部品質の要求事項を実現する際の観点の一例を示しているに過ぎない。したがって、読者は、自らが想定している機械学習利用システムの機能要件や構成等とユースケースを比較し、ユースケースの例示を柔軟に適用していくことが必要である。

3.3.1 配管の肉厚予測

※本事例は、「配管の肉厚予測」のイメージとして掲載するものであり、以降のユースケースの記述の前提ではない。

関連事例:「腐食による配管減肉量推定」(横河電機株式会社)⁵¹



石油精製で使用する常圧蒸留塔の上部配管は宿命的に腐食が発生する。腐食の進行を見える化することで、メンテナンスの効率化、運転調整による腐食の進行調節を行う。

蒸留塔上部配管に関連する 20 か所のプロセスデータと肉厚の定期測定値の 2 年分を学習データとし、プロセスデータと減肉進行の関係を回帰モデル(教師あり)でモデル化。オペレータの画面にリアルタイムの減肉の推定量を表示することで、メンテナンスのタイミングを適正化するとともに、オペレータが配管の状態をより意識した運転を実施することも可能になった。

1) ユースケースの前提

a. 概要

ユースケース「配管の肉厚予測」は、減肉の進行を予測に基づいて察知して安全を守ることと、逆に過度な点検や取り替え作業を削減してメンテナンスを効率化することを目的として、プロセスデータ等から現在の配管の肉厚を予測する機械学習利用システムである。

現状、減肉に伴う配管メンテナンスは主に定期的に行う時間基準保全(Time Based Maintenance: TBM)で実施されているが、定期的な点検間に急激な腐食が進行するケースや、まだメンテナンスの必要がない配管を点検・交換してしまう逸失利益が多く生じるケースがある。これを肉厚の予測に応じた状態基準保全(Condition Based Maintenance: CBM)に移行することで、保安力を向上するとともに逸失利益を削減することが期待できる。

⁵¹ 本事例は、「プラントにおける先進的 AI 事例集 ～AI プロジェクトの成果実現と課題突破の実践例～」に詳細を記載している。併せて参照されたい。

b. 機能要件

本ユースケースでは、機械学習利用システムの機能要件として、「配管の肉厚を予測する」を設定する。

c. 導入イメージ

本ユースケースでは、導入のイメージを図 3-2 のように設定している。従来のプラントの配管のメンテナンスは、設置された配管を決められた期間で定期的に保全員が確認している（TBM）⁵²。一方、機械学習利用システムの導入後は、機械学習要素が現在の配管の肉厚をリアルタイムで予測し、肉厚の予測値を保全員に出力する。保全員はこの出力内容と各種センサデータから、肉厚の実測を行うべきかどうか判断する。

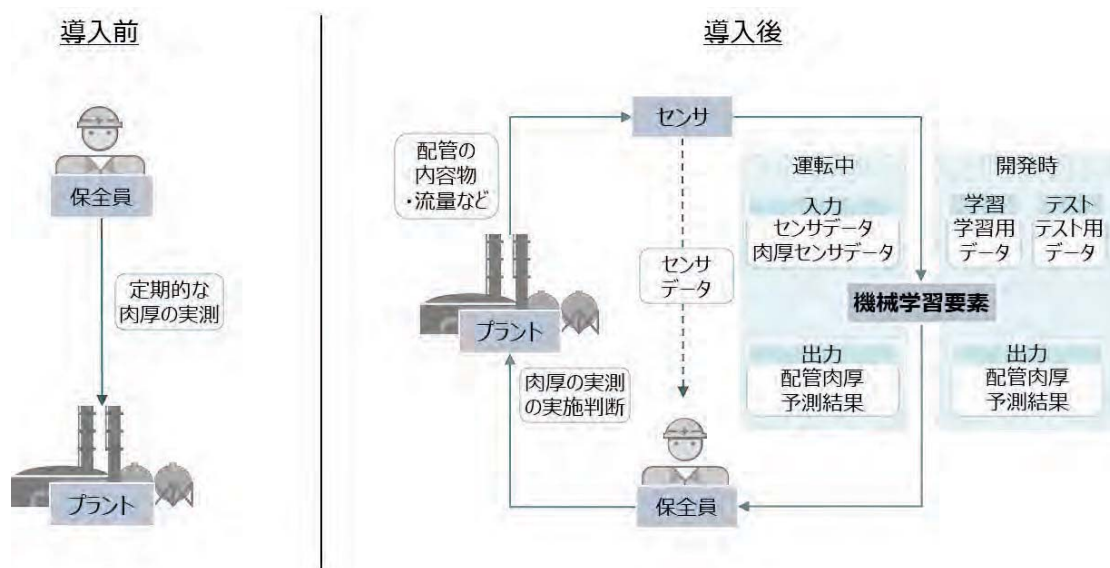


図 3-2 ユースケース「配管の肉厚予測」の導入イメージ（例）

d. 他のシステムとの関係

本ユースケースでは、機械学習要素と他のシステムとの関係を図 3-3 のように設定している。機械学習要素は、配管の内容物や流量などのデータを入力として、配管の肉厚を予測し、肉厚の予測値を出力し、肉厚予測値は保全員に対して提示される。保全員は肉厚予測値が、メンテナンスが必要なしきい値を超えているかを判断し、各種センサ等の従来収集しているデータも勘案して、メンテナンスが必要かどうかの判断を行う。

機械学習要素の出力（肉厚の予測値）を監視・補正する外部安全機構は存在しない。また、配管の内容物の漏洩を防ぐための機械学習利用システムと独立の安全関連系も存在しないことを想定する。

⁵² 現在、多くのプラントオーナー事業者は、配管のリスク等に応じて年に数回（年に1回～一ヶ月に1回程度）、作業員が肉厚の実測を行っている。なお、法定義務に関しては、例えば高圧ガス保安法およびその関連規則では、1年に1度の肉厚測定を求められており、それ以上の頻度は自主的な保安活動と位置づけられる。

このため、本ユースケースの前提の下では、「AISL 表」を参照する際に「③」の列を適用することとなる。

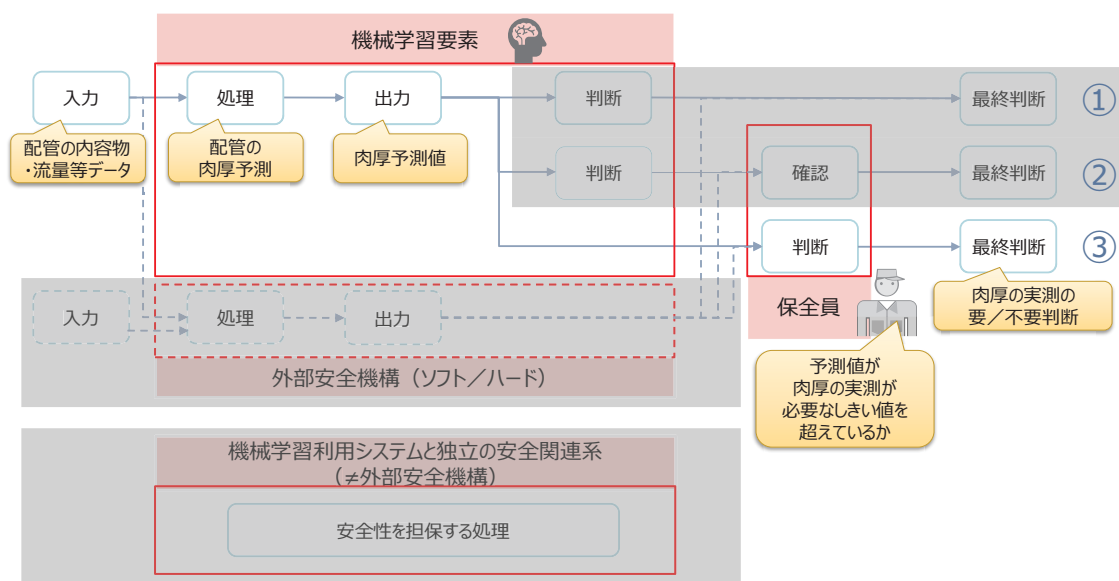


図 3-3 ユースケース「配管の肉厚予測」における機械学習要素と他のシステムの関係（例）

e. 機械学習要素の構成

本ユースケースでは、機械学習要素の構成を表 3-2 のように設定している。肉厚の値を予測するため、学習モデルとして、配管の減肉に影響すると考えられる内容物の種類・流量・流速・配管の材料・圧力などと、肉厚の値との関係を学習する教師ありの回帰モデルを想定する。学習データとしては配管の減肉に影響する各種データと肉厚の実測値を用い、テストデータにも同様のデータを用いる。

表 3-2 「配管の肉厚予測」の機械学習要素の構成（例）

学習手法	回帰(教師あり)
学習モデル	配管の内容物・流量・流速・材質・圧力などと肉厚値の関係を学習
運用時の入力データ	配管の内容物・流量・流速・材質・圧力データ
開発時の学習データ	配管の内容物・流量・流速・材質・圧力データ、肉厚センサデータ(実測値)
開発時のテストデータ	配管の内容物・流量・材質・流速・圧力データ、肉厚センサデータ(実測値)

2) 利用時品質・外部品質の項目例

本ユースケースの設定を前提とし、利用時品質・外部品質の項目を表 3-3 の通り設定する。2 章に示した階層的な品質確保の手順（図 3-4）においては、「①機械学習を用いたシステムで実現したいことを設定」「②機械学習要素に求める出力を設定し、その達成レベルを決める」に該当する。

表 3-3 ユースケース「配管の肉厚予測」の利用時品質・外部品質の項目例

利用時品質	外部品質
リスク回避性	
保全員による肉厚の実測が必要な配管を見落とさない(S-U1)	実際よりも肉厚を厚く予測する誤差を一定以内に抑える (S-E1)
パフォーマンス	
必要以上のメンテナンスを行わなくてよいようにする(P-U1)	実際よりも肉厚を薄く予測する誤差を一定以内に抑える(P-E1)

注) 各項目の記号は、項目間の関係を明らかにするために説明の都合上付与しているものであり、本ガイドラインとして付与を求めるものではなく、レベルの大小とも無関係である。なお、S: Safety、P: Performance、U: Use、E: External を意味する。

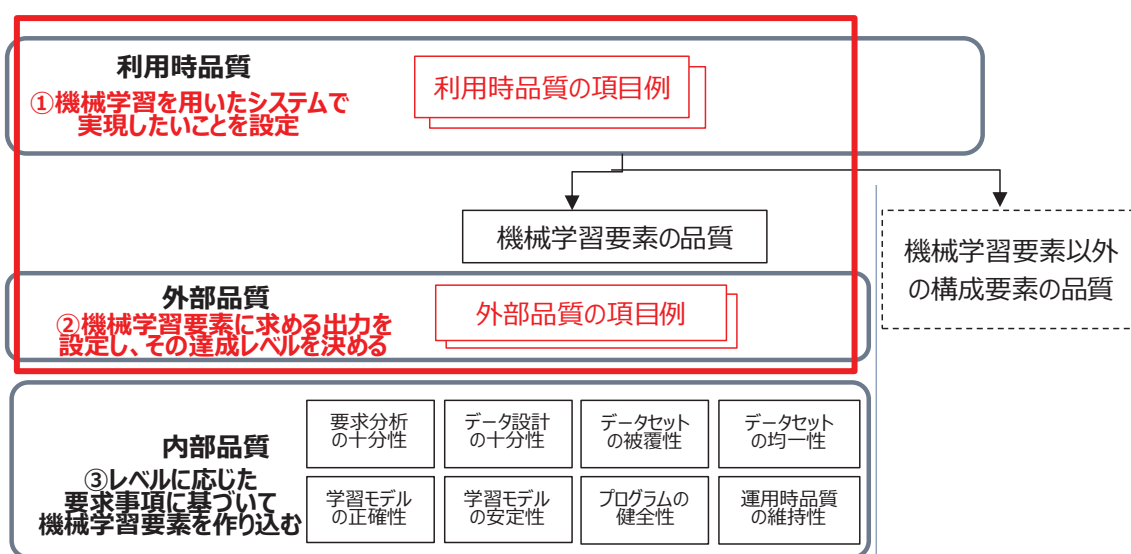


図 3-4 階層的な品質確保における利用時品質・外部品質項目設定の位置づけ

a. 「リスク回避性」に関する検討例

機能要件「配管の肉厚を予測する」を踏まえ、「リスク回避性」の観点から、機能要件が達成されないことで人的・経済的被害が生じることを防ぐための品質管理を行うために、管理対象である利用時品質及び外部品質を具体的に設定する。

● 利用時品質

人的・経済的被害の予防という観点からは、肉厚の予測を誤り、実際には減肉が進んで肉厚が要交換の基準値以下になってしまうということが望ましくない。このリスクを回避するために、「リスク回避性」属性の利用時品質として「保全員による肉厚の実測が必要な配管を見落とさない(S-U1)」という項目を設定する。

- 外部品質

設定した利用時品質に対し、機械学習要素のアウトプットに求める外部品質を定める。「保全員による肉厚の実測が必要な配管を見落とさない(S-U1)」という利用時品質に対応する外部品質は、「実際よりも肉厚を厚く予測する誤差を小さくする(S-E1)」と設定する。ここで「実際の肉厚との予測誤差」ではなく、「実際よりも肉厚を大きく予測する誤差」としているのは、単に予測が外れることをリスクととらえているのではなく、実際の肉厚よりも厚く予想してメンテナンスのタイミングが遅れる間に減肉が決定的に進み内容物が漏洩し、人的・経済的損失が発生するというリスクを想定しているためである。一方で「実際よりも肉厚を薄く予測する誤差」も考えられるが、こちらは後段で「パフォーマンス」軸の品質として設定する。

この外部品質に対し、必要な外部品質のレベル「AISL」を設定する。「実際よりも肉厚を厚く予測する誤差を小さくする(S-E1)」に対し、実際よりも肉厚を厚く予測した場合に想定される人的・経済的被害の大きさを検討し、基準に従って AISL を設定する。設定した AISL が機械学習要素の AISL となり、これに応じて内部品質の要求レベルが決まる。

b. 「パフォーマンス」に関する検討例

機能要件「配管の肉厚を予測する」を踏まえ、「パフォーマンス」の観点からは、機能要件を望ましい水準で達成するための品質管理を行うために、管理対象である利用時品質及び外部品質を具体的に設定する。

- 利用時品質

配管の肉厚予測では、作業員による肉厚の実測などのメンテナンスの頻度を適切な（＝必要以上に多くない）水準とすることが求められる。必要以上にメンテナンスの頻度が高くなると、保全のコストや機会損失が大きくなるため、「リスク回避性」軸の利用時品質で安全を確保することを前提として、可能な限り交換の頻度を抑えることが合理的である。そのため、利用時品質としては「必要以上のメンテナンスを行わなくてよいようにする(P-U1)」という項目を設定する。

- 外部品質

設定した利用時品質に対し、機械学習要素のアウトプットに求める外部品質を定める。「必要以上のメンテナンスを行わなくてよいようにする(P-U1)」という利用時品質に対応する外部品質は、「実際よりも肉厚を薄く予測する誤差を一定以内に抑える(P-E1)」と設定する。プラントの保全の管理者が認識する「必要以上のメンテナンスが発生すること」を、機械学習要素のアウトプットに置き換え、「実際よりも肉厚を薄く予測する」と表現している。

この外部品質に対し、必要な外部品質のレベル「AIPL」を設定する。「実際よりも肉厚を小さく予測する誤差を一定以内に抑える(P-E1)」に対し、求める正答率の水準や、必須とする度合いを検討し、基準に従って AIPL を設定する。なお、配管のメンテナンスについては交換について、（「リスク回避性」で安全を確保することは前提に）経済性をできるだけ追求する厳しい要求水準を求める場合、必要に応じて「AIPL2」を適用することも考

えられる。

3) 内部品質確保のための「ユースケース固有の観点」

本ユースケースの設定を前提とした場合の、内部品質の各要求事項を実現するための留意点（「観点」）を表 3-4 に示す。以下に記載する「観点」は、本ユースケースに近い機械学習要素を開発する場合に、参考にすることができる。2 章に示した階層的な品質確保の手順（図 3-5）においては、「③レベルに応じた要求事項に基づいて機械学習要素を作り込む」に該当する。

表 3-4 「配管の肉厚予測」に係る「ユースケース固有の観点」

※本ケースの内部品質の要求事項と観点の一覧は附録のチェックリストを参照

内部品質	要求事項 ⁵³	ユースケース固有の観点
要求分析の 十分性	(要求事項共通)	腐食の種類によって「データ設計の十分性」「データセットの被覆性」の評価にも影響するため、どの腐食を対象とするかという点まで絞り込む。
データ設計 の十分性	(Lv1)また、特に重要と考えられる環境要因の差異に対する属性を抽出し、大きなリスクの要因との組み合わせに対応するケースを用意すること。	「環境要因」とは、ここでは気候、塩分(海からの距離や風向などの地域性)などを指す。
データセット の被覆性	(Lv1)テスト用データセットの取得源や方法を検討し、応用の状況に対して偏りが無いことを期待できるようにすること。	「応用の状況」とは、ここでは対象とする配管や観測の頻度、評価の時間軸(リアルタイムの予測を行うか等)などを指す。
	(要求事項共通)	配管の内容物・流量・材質・流速・圧力など、前提としている属性のデータの範囲がカバーできているか注意する。
データセット の均一性	(要求事項共通)	- 上記の属性のカバーすべきデータの各範囲のデータ量が十分であることに注意する。 - ある範囲のデータ量が十分でない場合、その範囲の予測精度が低くなることを認識する。
機械学習モ デルの正確 性	—	—
機械学習モ デルの安定 性	—	—

⁵³ 表中では、「機械学習品質マネジメントガイドライン 第1版」記載の要求事項について、「ユースケース固有の観点」に関連するもののみ抜粋している。本表に記載していない項目も、要求事項に含まれる。

内部品質	要求事項 ⁵³	ユースケース固有の観点
プログラムの健全性	—	—
運用時品質の維持性	(要求事項共通)	- 既存手法による交換可否判断や、実際の交換時の配管の状態等を踏まえ、実際の正解度や見落としの有無を検証する。 - 本ケースは特に精度の維持の重要性が高いため、モデル構築時の前提、教師データ類を整理しておき、運用開始後の変更管理において都度確認する。

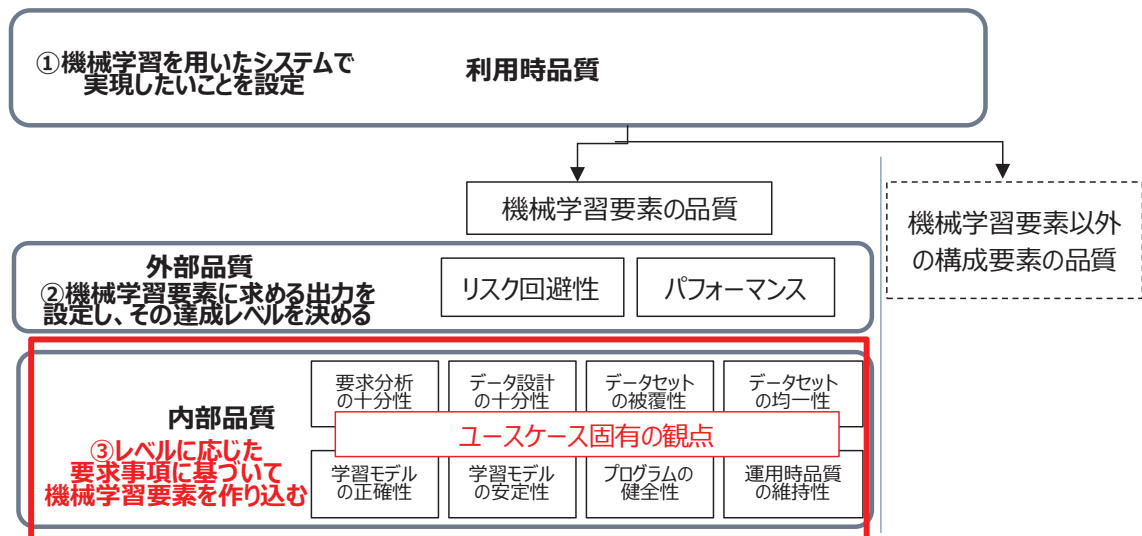


図 3-5 階層的な品質確保における「ユースケース固有の観点」の位置づけ

3.3.2 配管の画像診断

※本事例は、「配管の画像診断」のイメージとして掲載するものであり、以降のユースケースの記述の前提ではない。

関連事例:「画像の自動判別による配管外面の腐食箇所の抽出」
(プラントオーナー企業 X 社)



配管外面の腐食の検出は、定期的な点検作業によって行われており、作業負荷が高い。特に高所の配管は足場が必要になり、高頻度の点検は難しい。そこで、本システムでは、機械学習による腐食箇所のスクリーニング技術と、ドローンによる配管画像撮影技術を組み合わせることで、点検作業の負荷削減、点検範囲の拡大、点検の高頻度化を実現する。

過去に撮影した配管外面の画像データ(腐食無し／腐食有りのラベル付き)を用いた分類モデル(教師あり)を利用。撮影した配管外面の画像をインプットとし、腐食有りと判定した画像をスクリーニング結果として保全員に提示する。保全員が画像を確認した場合、現場で腐食状況を確認し、対応要否を判断する。

1) ユースケースの前提

a. 概要

ユースケース「配管の画像診断」は、配管外面の点検作業の負荷軽減を目的として、保全員による目視点検が必要な箇所をスクリーニングする機械学習利用システムである。従来、配管外面の腐食の点検は定期的に保全員が目視で実施しており、大きな労力を要している。スクリーニングにより保全員が目視で点検すべき箇所を限定することができれば、労力削減効果が大きい。また、ベテラン保全員に依存する業務量を減らすことができ、中長期的に保安力を維持することができる。

本ユースケースは画像の撮影手段を限定するものではないが、上記の関連事例のようにドローンによる撮影と組み合わせて活用する場合は、高所作業用の足場が必要になる高所の配管の点検が高頻度でできるため、更に大きなコスト削減効果と安全性の向上が期待できる。

b. 機能要件

本ユースケースでは、機械学習利用システムの機能要件として、「配管の画像から目視点検要否を判定する」を設定する。

c. 導入イメージ

本ユースケースでは、導入のイメージを図 3-6 のように設定している。従来の配管の腐食有無の判断は、定期的に保全員が配管を目視で確認することで行っていた。一方、機械学習利用システム導入後は、配管の画像から、当該箇所の腐食が保全員による目視点検が必要か否かを機械学習要素が判定し、保全員に対して要否判定結果を出力する。保全員は当該箇所の画像データを確認し、現場での目視点検結果によって配管の補修や交換を実施すべきかを判断する。

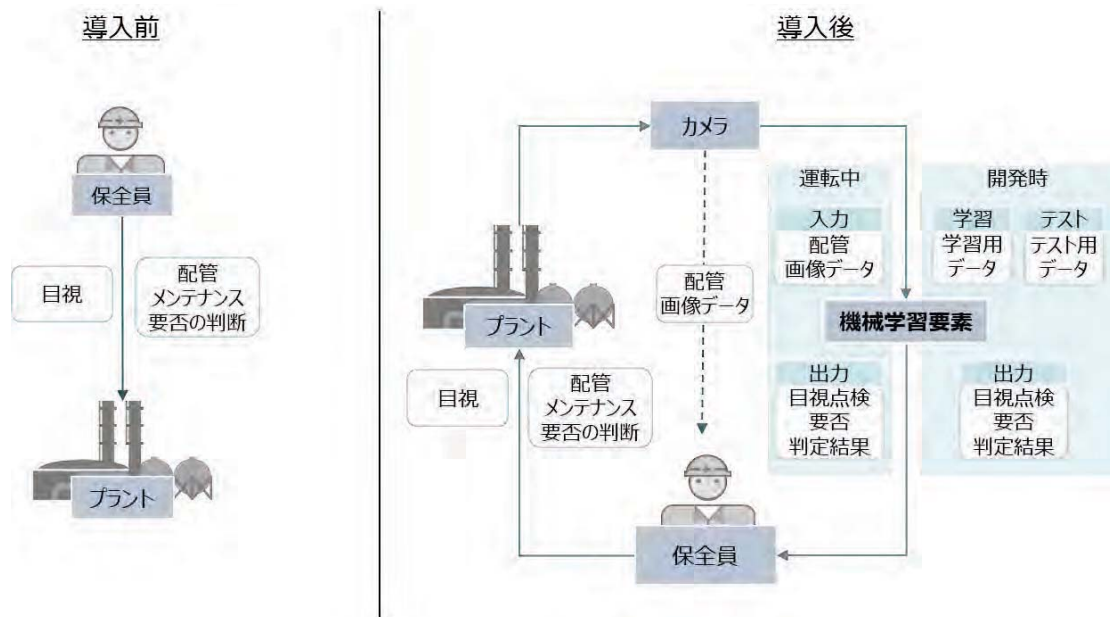


図 3-6 ユースケース「配管の画像診断」の導入イメージ（例）

d. 他のシステムとの関係

本ユースケースでは、機械学習要素と他のシステムとの関係を図 3-7 のように設定している。機械学習要素は、配管画像データを入力として目視点検の要否を判定し、要否の判定結果を保安員に対して出力する。本ユースケースの機械学習要素の出力は、目視点検の実施要否のスクリーニングを目的としたものであるため、判断の内容によって後段のフローが分岐すると想定する。目視点検が不要であると判断した場合、目視点検は実施しないという最終判断結果となる。目視点検が必要であると判断した場合、保安員が画像を確認して本当に目視点検が必要かどうかを判断し、その判断をもって最終判断とすると想定する。機械学習要素の出力（目視点検の要否）そのものを監視・補正する外部安全機構は存在しないと考ええる。このため、「AISL 表」を参照する際には、より高いレベルが設定される①の列を適用する必要がある。また、配管の内容物の漏洩を防ぐための機械学習利用システムと独立の安全関連系も存在しないことを想定する。

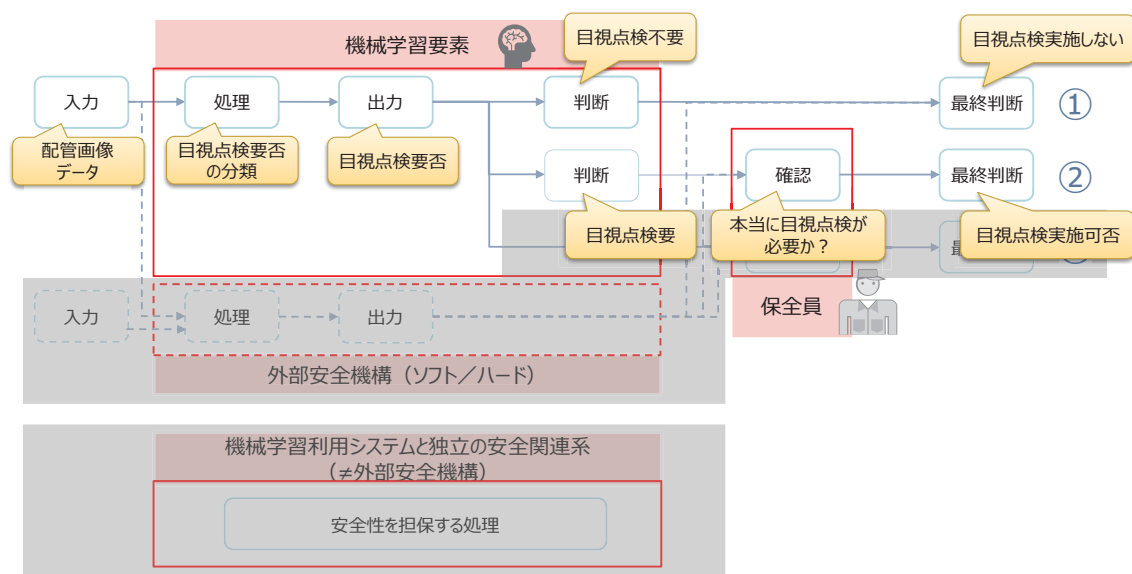


図 3-7 ユースケース「配管の画像診断」における
機械学習要素と他のシステムの関係（例）

e. 機械学習要素の構成

本ユースケースでは、機械学習要素の構成を表 3-5 のように設定している。過去の定期的な配管の保全活動等において配管を撮影した画像が収集されており、それらの画像の腐食の状態に対して目視点検の要否のラベル付けを行う。学習モデルとしては、それらの画像データを背景として、配管画像の特徴から目視点検要否を分類する教師ありの分類モデルを想定する。学習データとしては配管画像データとその配管画像の目視点検要否ラベルを用い、テストデータにも同様のデータを用いる。

表 3-5 「配管の画像診断」の機械学習要素の構成（例）

学習手法	分類(教師あり)
学習モデル	配管画像の特徴から目視点検要否を分類する分類モデル
運用時の入力データ	配管画像データ
開発時の学習データ	配管画像データ+目視点検要否ラベル
開発時のテストデータ	配管画像データ+目視点検要否ラベル

2) 利用時品質・外部品質の項目例

本ユースケースの設定を前提とし、利用時品質・外部品質の項目を表 3-6 の通り設定する。2 章に示した階層的な品質確保の手順（図 3-8）においては、「①機械学習を用いたシステムで実現したいことを設定」「②機械学習要素に求める出力を設定し、その達成レベルを決める」に該当する。

表 3-6 ユースケース「配管の画像診断」の利用時品質・外部品質の項目例

利用時品質	外部品質
リスク回避性	
目視点検が必要な配管を見落とさない(S-U1)	目視点検が「要」である場合に、「不要」であると判定する誤判定率を限りなく小さくする(S-E1)
パフォーマンス	
保全員の目視点検実施数を少なく抑える(P-U1)	目視点検が「不要」である場合に、「要」であると判定する誤判定率を一定以内にする(P-E1)

注) 各項目の記号は、項目間の関係を明らかにするために説明の都合上付与しているものであり、本ガイドラインとして付与を求めるものではなく、レベルの大小とも無関係である。なお、S: Safety、P: Performance、U: Use、E: External を意味する。

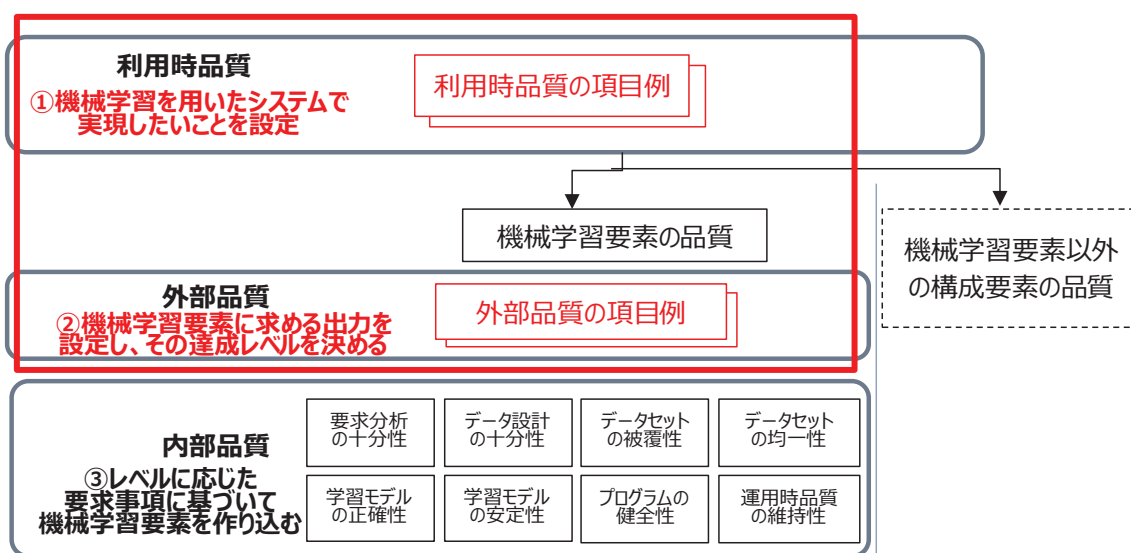


図 3-8 階層的な品質確保における利用時品質・外部品質項目設定の位置づけ

a. 「リスク回避性」に関する検討例

機能要件「配管の画像から目視点検要否を判定する」を踏まえ、「リスク回避性」の観点からは、機能要件が達成されないことで人的・経済的被害が生じることを防ぐための品質管理を行うために、管理対象である利用時品質及び外部品質を具体的に設定する。

● 利用時品質

人的・経済的被害の予防という観点からは、目視点検が必要な配管が見逃されることは望ましくない。このリスクを回避するために、「リスク回避性」属性の利用時品質として「目視点検が必要な配管を見落とさない(S-U1)」という項目を設定する。

● 外部品質

設定した利用時品質に対し、機械学習要素のアウトプットに求める外部品質を定める。

「目視点検が必要な配管を見落とさない(S-U1)」という利用時品質に対応する外部品質は、「目視点検が『要』である場合に、『不要』であると判定する誤判定率を限りなく小さくする(S-E1)」と設定する。

この外部品質に対し、必要な外部品質のレベル「AISL」を設定する。「目視点検が『要』である場合に、『不要』であると判定する誤判定率を限りなく小さくする(S-E1)」に対し、目視点検が必要な配管画像を誤って不要であると判定した場合に想定される人的・経済的被害の大きさを検討し、基準に従って AISL を設定する。なお、機械学習要素が「目視点検不要」と判断する場合、保全員が目視確認を行うことが無いため、AISL 表において、「人間による代替システムが用意されておらず、機械学習要素の判断結果がそのまま運転・点検に反映される」基準に相当するとして AISL の設定を行う。設定した AISL が機械学習要素の AISL となり、これに応じて内部品質の要求レベルが決まる。

b. 「パフォーマンス」に関する検討例

機能要件「配管の画像から目視点検要否を判定する」を踏まえ、「パフォーマンス」の観点からは、機能要件を望ましい水準で達成するための品質管理を行うために、管理対象である利用時品質及び外部品質を具体的に設定する。

● 利用時品質

配管の画像診断では、目視点検の頻度を適切な水準とすることが求められる。必要以上に目視点検の頻度が高くなると、労力の削減の効果が得られず、望ましくない。そのため、利用時品質としては「保全員の目視点検実施数を少なく抑える(P-U1)」という項目を設定する。

● 外部品質

設定した利用時品質に対し、機械学習要素のアウトプットに求める外部品質を定める。「保全員の目視点検実施数を少なく抑える(P-U1)」という利用時品質に対応する外部品質は、「目視点検が『不要』である場合に、『要』であると判定する誤判定率を一定以内にする(P-E1)」と設定する。プラントの保全の管理者が認識する「必要以上の目視点検が発生すること」を、機械学習要素のアウトプットに置き換え、「実際には目視点検が不要である配管画像を、目視点検が必要と判定する誤判定」と表現している。

この外部品質に対し、必要な外部品質のレベル「AIPL」を設定する。「目視点検が『不要』である場合に、『要』であると判定する誤判定率を一定以内にする(P-E1)」に対し、求める正答率の水準や、必須とする度合いを検討し、基準に従って AIPL を設定する。

3) 内部品質確保のための「ユースケース固有の観点」

本ユースケースの設定を前提とした場合の、内部品質の各要求事項を実現するための留意点（「観点」）を表 3-7 に示す。以下に記載する「観点」は、本ユースケースに近い機械学習要素を開発する場合に、参考にすることができる。2 章に示した階層的な品質確保の手順（図 3-9）においては、「③レベルに応じた要求事項に基づいて機械学習要素を作り込む」に該当する。

表 3-7 「配管の画像診断」に係る「ユースケース固有の観点」

※本ケースの内部品質の要求事項と観点の一覧は附録のチェックリストを参照

内部品質	要求事項 ⁵⁴	ユースケース固有の観点
要求分析の 十分性	(要求事項共通)	・ 保温材が巻かれている配管を取り扱う場合は、配管そのものの劣化ではなく、保温材の劣化が対象となる点に注意する。
データ設計 の十分性	・ (Lv1)また、特に重要と考えられる環境要因の差異に対する属性を抽出し、大きなリスクの要因との組み合わせに対応するケースを用意すること。	・ 「環境要因」とは、ここでは日照、天気、季節、時間帯などを指す。
	(要求事項共通)	・ 配管によって塗装や防錆塗装あるいは配管そのものの地肌の色が異なっているため、それらの違いを加味して精度を確保する。 ・ 配管上に積雪があるケース等、画像により直接配管外面を確認することができない場合があることを把握する。 ・ 撮影方法のルール・留意点の整備によりデータの品質を一定レベルとすることについて検討する。 ・ 画像のブレに対処するためモデルにより吸収することも考えられるが、システムの複雑化・不確定要素の増加を招く可能性を把握する。
データセット の被覆性	(要求事項共通)	・ 周囲の環境(日照、時刻など)やドローン撮影での画像のブレなど、入力画像が不鮮明になる場合の対応を検討する。 ・ 環境要因の各属性のデータの範囲がカバーできているか注意する。
データセット の均一性	(要求事項共通)	・ 上記の属性のカバーすべきデータの各範囲のデータ量が十分であることに注意する。 ・ ある範囲のデータ量が十分でない場合、その範囲の予測精度が低くなることを認識しておく。
機械学習モデルの正確性	—	—
機械学習モデルの安定性	—	—

⁵⁴ 表中では、「機械学習品質マネジメントガイドライン 第1版」記載の要求事項について、「ユースケース固有の観点」に関連するもののみ抜粋している。本表に記載していない項目も、要求事項に含まれる。

内部品質	要求事項 ⁵⁴	ユースケース固有の観点
プログラムの健全性	—	—
運用時品質の維持性	(要求事項共通)	運用段階で撮影した画像での精度検証を行う。「目視点検要」と判定され実際に目視点検を実施した結果を記録し、これと対照して、判定精度が低い場合は、判定に用いた入力画像やモデルを重点的にチェックする。

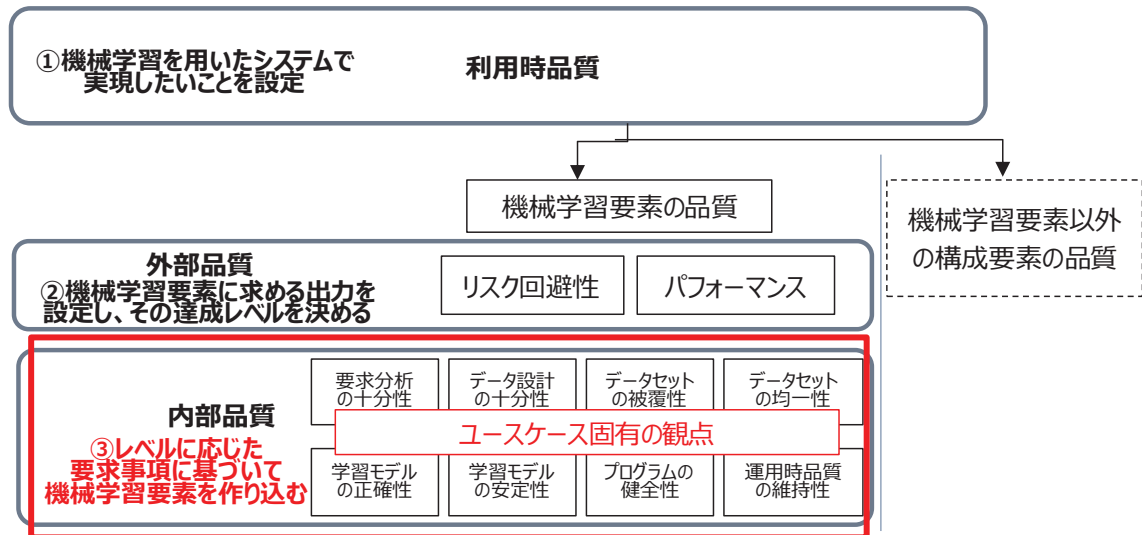
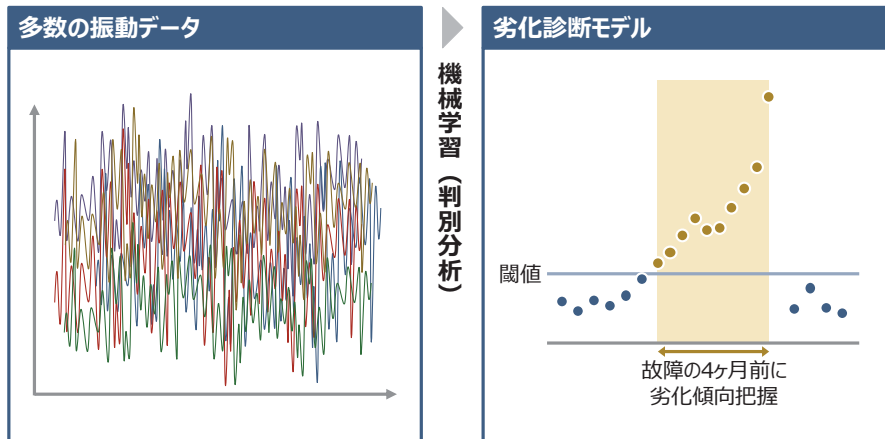


図 3-9 階層的な品質確保における「ユースケース固有の観点」の位置づけ

3.3.3 設備劣化診断

※本事例は、「設備劣化診断」のイメージとして掲載するものであり、以降のユースケースの記述の前提ではない。

関連事例:「攪拌機の劣化診断」(プラントオーナー企業 Y 社)



※ 劣化傾向の早期把握により、メンテナンス計画や調達を最適化する目的で導入 (安全関連系とは独立)

これまで、攪拌機⁵⁵の各部に設置された振動センサにより、故障の数週間前に異常な振動を検出し、設備の停止判断やメンテナンスを実施していた。

しかし、機械学習で振動データ等を分析すると、攪拌機の軸受けと減速機の振動変位の間に存在する通常の相関関係が、故障の数か月前から逸脱することが判明した。数か月前から劣化の兆候を検出できるようになれば、メンテナンス計画や調達の最適化が期待できる。

過去の振動データに対し、軸受けと減速機の振動変位の相関を変数化し、上記した通常の相関関係からの逸脱度をマハラノビス距離⁵⁶を用いた判別分析により指標化し危険度として出力する。なお、本システムは数ヶ月単位のメンテナンス・調達計画の最適化のみに活用しており、設備故障の把握による安全確保は、本システムと独立のシステムで別途実施している。

1) ユースケースの前提

a. 概要

ユースケース「設備劣化診断」は、プラント設備の特定の部品の劣化傾向の早期発見を目的として、プラント設備の数週間～数か月の長期的な未来に顕在化する異常の予兆を事前に検知する機械学習利用システムである。プラント設備の部品劣化を早い段階で把握できれば、メンテナンス計画や部材調達計画への反映、劣化を遅らせる運転の実施など、プラントの保全活動において効果が大きい。

なお、数日～数十分後の短期的な未来に発生し得るプラントの異常を検知する機械学

⁵⁵ タンク内で液体と液体、液体と個体等を混ぜ合わせ、化学反応や混合等を進めるための装置。

⁵⁶ データの分散を考慮して、あるデータがデータ群からどの程度離れているかを表す量。異常データの判別に用いられることが多い。

習利用システムのユースケースは「3.3.4 異常予兆検知・診断」を参照。

b. 機能要件

本ユースケースでは、機械学習利用システムの機能要件として、「個別設備の部品の将来の劣化傾向を予測する」を設定する。

c. 導入イメージ

本ユースケースでは、導入のイメージを図 3-10 のように設定している。従来のプラント設備の劣化診断は、プラント設備に設置された複数のセンサの値を定期的に保全員が読み取り、センサの値が予め定められたしきい値から逸脱していた場合に交換実施の判断を行っていた。一方、機械学習利用システムの導入後は、機械学習要素が設備の劣化傾向有無を判定し、劣化傾向有りと判定された場合に保全員に対してアラートを提示する。保全員はこの提示内容と設備の稼働データから、部品の点検や交換の要否を判断する。

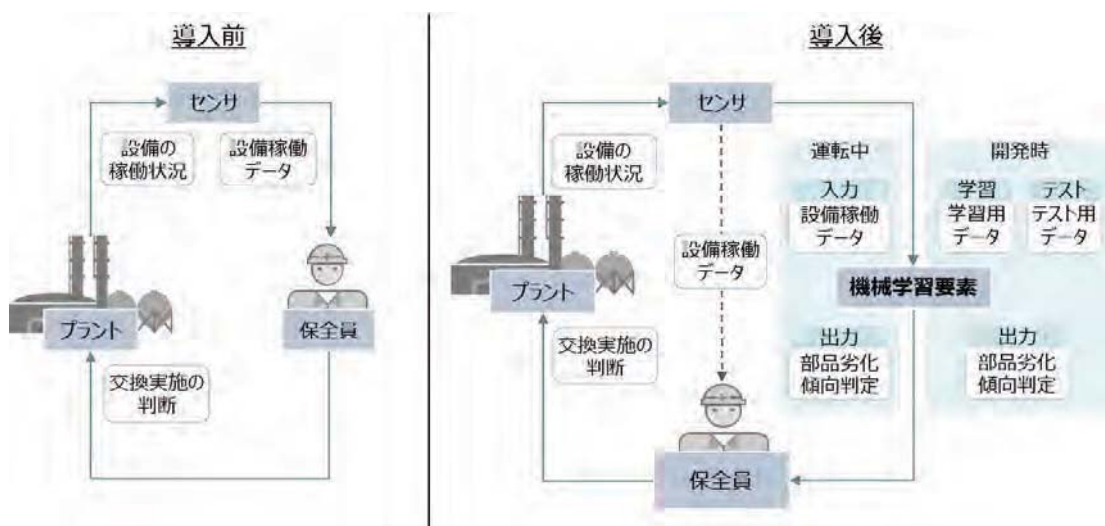


図 3-10 ユースケース「設備劣化診断」の導入イメージ（例）

d. 他のシステムとの関係

本ユースケースでは、機械学習要素と他のシステムとの関係を図 3-11 のように設定している。機械学習要素は、設備の稼働データを入力として劣化傾向の有無を分類し、その結果を保全員に出力する。保全員は、設備の対象部材の劣化傾向が交換・メンテナンスが必要なレベルかどうかを判断し、最終判断として交換・メンテナンスの要否の判断を行う。このため、本ユースケースの前提の下では、「AISL 表」を参照する際に「③」の列を適用することとなる。

機械学習要素の判断（設備の劣化傾向の有無）そのものを監視・補正する外部安全機構は存在しないとする。

なお、本ユースケースの機械学習利用システムは、中長期的に故障の兆候を検知して保全活動を効率化するために設備の劣化を診断するものとし、差し迫った設備の故障状況

を検知する既存システムは全く別の系として存在するものとの前提を置く。

このように、機械学習利用システムは純粋に保全活動の効率化を目的としたもので、安全性は既存システムで担保するものと想定する。そのため、本ユースケースの前提の下では、「リスク回避性」の評価は実施しない。これは、あくまで本ガイドラインで例示した前提のもとで判断されるものであり、本ユースケースに類似するあらゆるケースで「リスク回避性」の検討が不要であることを意味しない。機械学習要素に求める安全機能の水準は、個別の適用ごとに判断する必要がある。

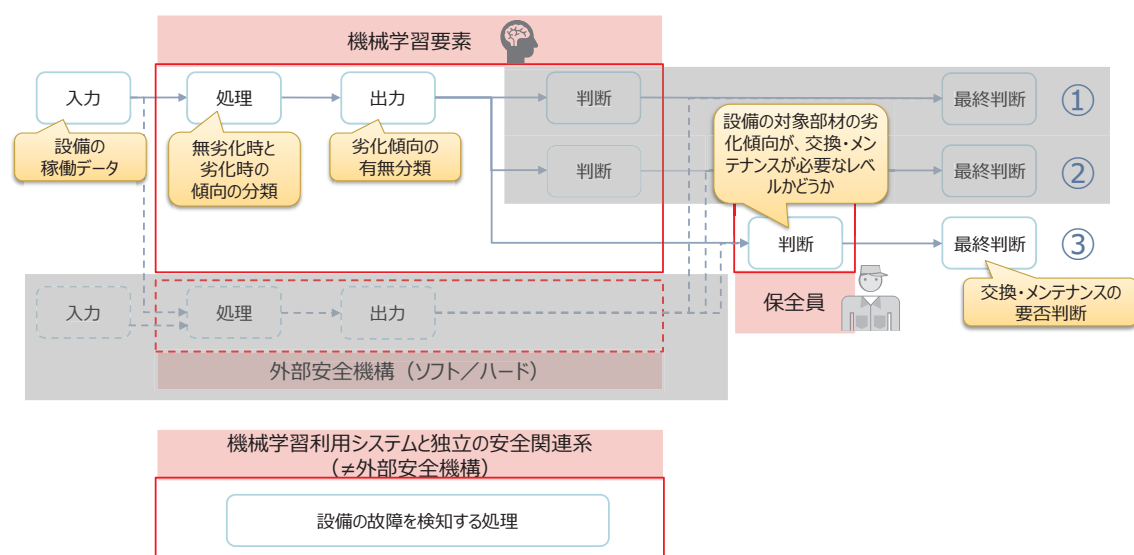


図 3-11 ユースケース「設備劣化診断」における機械学習要素と他のシステムの関係（例）

e. 機械学習要素の構成

本ユースケースでは、機械学習要素の構成を表 3-8 のように設定している。学習モデルとして、劣化がある場合と無い場合のそれぞれの変数の分布を学習する判別分析の分類モデルを想定する。したがって、ある時点の設備の稼働時系列データを入力とした場合、稼働データが劣化時の分布に分類された場合に劣化傾向有りとして出力する。学習データとしては過去の設備の稼働時系列データと将来の劣化有無ラベルを用い、テストデータにも同様のデータとラベルを用いる。劣化有無のラベル付与はその時点ではなく将来の劣化発生有無のラベルを付与することを想定する。ここでの「将来」は数週間～数か月の長期的な未来を想定する。

表 3-8 「設備劣化診断」の機械学習要素の構成（例）

学習手法	分類(教師あり)
学習モデル	劣化時・無劣化時の稼働データから、変数の分布を学習
運用時の入力データ	設備稼働時系列データ
開発時の学習データ	過去の設備稼働時系列データ+将来の劣化有無ラベル
開発時のテストデータ	過去の設備稼働時系列データ+将来の劣化有無ラベル

2) 利用時品質・外部品質の項目例

本ユースケースの設定を前提とし、利用時品質・外部品質の項目を表 3-9 の通り設定する。本ユースケースでは、機械学習利用システムは純粋に保全活動の効率化を目的としたもので、安全性は既存システムで担保するものとする。そのため、「リスク回避性」の利用時品質・外部品質は設定せず、AISL の評価は実施しない。2 章に示した階層的な品質確保の手順（図 3-12）においては、「①機械学習を用いたシステムで実現したいことを設定」「②機械学習要素に求める出力を設定し、その達成レベルを決める」に該当する。

表 3-9 ユースケース「設備劣化診断」の利用時品質・外部品質の項目例

利用時品質	外部品質
リスク回避性	
—	—
パフォーマンス	
部品の劣化を正しく診断する(P-U1)	「劣化あり」「劣化なし」の分類誤差を一定以内に抑える(P-E1)
メンテナンス計画への反映が可能な程度に十分に早いタイミングに劣化の傾向を予測する(P-U2)	「劣化なし」から「劣化あり」に変化したとの判定結果を、決められた時間前に出力する(P-E2)

注) 各項目の記号は、項目間の関係を明らかにするために説明の都合上付与しているものであり、本ガイドラインとして付与を求めるものではなく、レベルの大小とも無関係である。なお、S: Safety、P: Performance、U: Use、E: External を意味する。

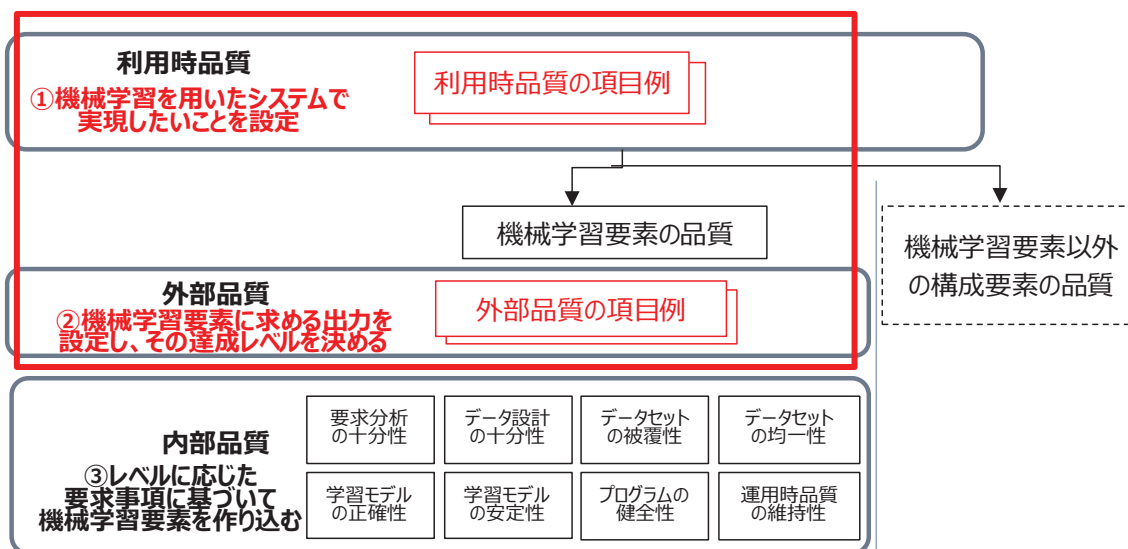


図 3-12 階層的な品質確保における利用時品質・外部品質項目設定の位置づけ

a. 「リスク回避性」に関する検討例

本ユースケースでは「リスク回避性」の利用時品質・外部品質を設定しないため、AISL の評価は実施しない。

b. 「パフォーマンス」に関する検討例

機能要件「個別設備の部品の将来の劣化傾向を予測する」を踏まえ、「パフォーマンス」の観点からは、機能要件を望ましい水準で達成するための品質管理を行うために、管理対象である利用時品質及び外部品質を具体的に設定する。

● 利用時品質

設備劣化診断では、劣化診断の精度や劣化傾向の検出タイミングを適切な水準とすることが求められる。劣化診断の精度水準を定める利用時品質として、「部品の劣化を正しく診断する(P-U1)」を設定する。劣化傾向の検出タイミングについては、数週間～数か月間単位でのメンテナンスの計画に反映できるタイミングで将来の劣化傾向を検出する必要がある。そのため、利用時品質としては「メンテナンス計画への反映が可能な程度に十分に早いタイミングに劣化の傾向を予測する(P-U2)」という設定となる。

● 外部品質

それぞれの利用時品質に対し、機械学習要素のアウトプットに求める外部品質を定める。「部品の劣化を正しく診断する(P-U1)」という利用時品質に対応する外部品質は、「『劣化あり』『劣化なし』の分類誤差を一定以内に抑える(P-E1)」と設定する。また、「メンテナンス計画への反映が可能な程度に十分に早いタイミングに劣化の傾向を予測する(P-U2)」という利用時品質に対応する外部品質は、「『劣化なし』から『劣化あり』に変化したとの判定結果を、決められた時間前に出力する(P-E1)」と設定する。

それぞれの外部品質に対し、必要な外部品質のレベル「AIPL」を設定する。「『劣化あり』『劣化なし』の分類誤差を一定以内に抑える(P-E1)」については、「分類誤差を目標値に抑える」ことをどの程度求めるかの水準や、必須とする度合いを検討し、基準に従って AIPL を設定する。同様に、「『劣化なし』から『劣化あり』に変化したとの判定結果を、決められた時間前に出力する(P-E2)」については、求める出力タイミングの水準や、必須とする度合いを検討し、基準に従って AIPL を設定する。これらのうち最も大きい AIPL が、機械学習要素の AIPL となり、これに応じて内部品質の要求レベルが決まる。

3) 内部品質確保のための「ユースケース固有の観点」

本ユースケースの設定を前提とした場合の、内部品質の各要求事項を実現するための留意点（「観点」）を表 3-10 に示す。以下に記載する「観点」は、本ユースケースに近い機械学習要素を開発する場合に、参考にすることができる。2 章に示した階層的な品質確保の手順（図 3-13）においては、「③レベルに応じた要求事項に基づいて機械学習要素を作り込む」に該当する。

表 3-10 「設備劣化診断」に係る「ユースケース固有の観点」

※本ケースの内部品質の要求事項と観点の一覧は附録のチェックリストを参照

内部品質	要求事項 ⁵⁷	ユースケース固有の観点
要求分析の 十分性	(要求事項共通)	・ 製造条件によって変化する製品の成分値に対して、どの成分値の範囲までを対象とするか検討する。製造する製品が異なる場合だけでなく、流体(※)やプロセスが変わった場合も含む。 (※)混相流・多相流の配分変化など
データ設計 の十分性	・ (Lv1)また、特に重要と考えられる環境要因の差異に対する属性を抽出し、大きなリスクの要因との組み合わせに対応するケースを用意すること。	・ 「環境要因」とは、ここでは設置場所、動作環境、気温・湿度、運転方法、原材料、用役などを指す。
	(要求事項共通)	・ 対象とする製品の成分値の範囲に対して、それらの学習データが収集できるかどうか検討する ・ シミュレータのデータを利用する場合、環境要因の変化(例:高湿度→低湿度)がシミュレータで考慮されているか確認する。 ・ データセットをシミュレーションで取得する場合は、シミュレータの妥当性を十分検証する。 ・ 部材の変更直後は基本的に「劣化無し」のデータとしてよい。「劣化なし」期間は部材のスペックに依るが、使用環境で変化することに注意する(過去の交換頻度等を参照して、「劣化なし」とする期間を決定する)。 ・ 部材の変更直後で「慣らし運転⁵⁸」が必要な場合は、その期間はデータを収集しないなどの管理を行う。
データセット の被覆性	・ (Lv1)テスト用データセットの取得源や方法を検討し、応用の状況に対して偏りが無いことを期待できるようにすること。	・ 「応用の状況」とは、ここでは対象とする設備の種類や稼働状況(常時／一時、負荷の変更等)などを指す。
	(要求事項共通)	・ 稼働データの劣化有無ラベルの確認は、その判断が適切に行える専門性を持った要員によって行う。

⁵⁷ 表中では、「機械学習品質マネジメントガイドライン 第1版」記載の要求事項について、「ユースケース固有の観点」に関連するもののみ抜粋している。本表に記載していない項目も、要求事項に含まれる。

⁵⁸ 部材の変更直後、本格的な生産の前に低負荷で稼働する等により設備の状態に問題がないか確認する運転のこと。

内部品質	要求事項 ⁵⁷	ユースケース固有の観点
データセットの均一性	(要求事項共通)	「劣化なし」として想定する様々な状態の稼働データが偏りなく取得できていること。 - ある状態の稼働データが十分取得できない場合、その状態から逸脱する劣化を検出する精度が低下する可能性があることを認識すること。
機械学習モデルの正確性	—	—
機械学習モデルの安定性	—	—
プログラムの健全性	—	—
運用時品質の維持性	(要求事項共通)	- 部材の種類を変えた場合には、再学習、学習モデルの切り替えなどの対応が必要になる可能性があることに留意する。 - 変化する製品成分値の範囲に対して、製品成分値の入力時に品質が維持されているか注意する。 - 対象設備そのものだけでなく、周辺の状態を含め、当初想定していた環境要因などの前提条件から逸脱していないか確認する。

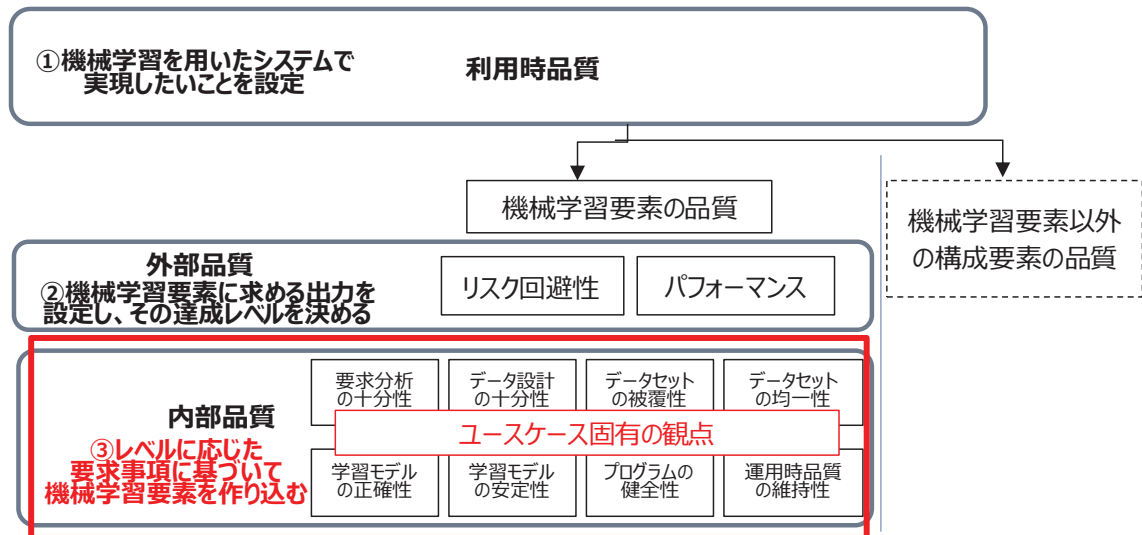
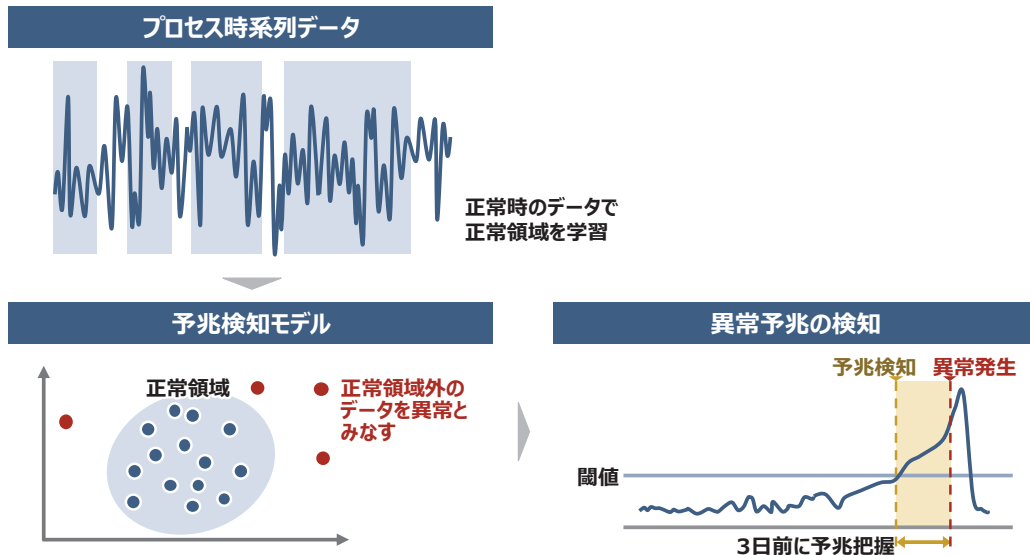


図 3-13 階層的な品質確保における「ユースケース固有の観点」の位置づけ

3.3.4 異常予兆検知・診断

※本事例は、「異常予兆検知・診断」のイメージとして掲載するものであり、以降のユースケースの記述の前提ではない。

関連事例:「動力プラントにおける異常予兆検知・診断」(プラントオーナー企業 Z 社)



動力部では、供給支障が無いよう常に各設備の運転値の上下限監視をしているが、上下限監視で捉えきれない異常が要因で、設備が突発停止することがある。そこで、本システムにより、設備の異常兆候を早期に検知し、設備の突発停止を削減する。

正常時の時系列データを用いた外れ値検出モデル(教師なし)を利用。DCSより取得したプロセス運転データをインプットとし、正常時からの逸脱の大きさを数値化した異常度のグラフや、異常度に影響を与えている変数を表示する。

1) ユースケースの前提

a. 概要

ユースケース「異常予兆検知・診断」は、異常によるプラントの突発停止の回避を目的として、プラントの数日～数十分後の短期的な未来に顕在化し得る異常の予兆を事前に検知する機械学習利用システムである。突発停止が発生すると、たとえ事故につながらない場合であっても、再稼働までに大きなコストを要する。このため、事前に異常の予兆を検知することにより、通常手順での停止や、異常発生箇所への対応が可能になることは、プラントにとって安全性・生産性の両面で効果が大きい。なお、長期的な設備の劣化状況进行分析する機械学習要素のケースは、「3.3.3 設備劣化診断」を参照。

b. 機能要件

本ユースケースでは、機械学習利用システムの機能要件として、異常予兆の検知に相当

する「異常予兆検知時に発報する」、異常内容の診断に相当する「異常箇所、異常の深刻度、異常と関連のある変数を出力する」を設定する。

c. 導入イメージ

本ユースケースでは、導入のイメージを図 3-14 のように設定している。従来のプラントの異常検知は、プラントに設置された複数のセンサの値を 1 日に数回運転員が読み取り、運転員の経験から将来の異常の発生有無を判断していた。一方、機械学習利用システムの導入後は、機械学習要素が将来の異常発生有無（異常予兆の有無）をリアルタイムで監視し、異常の予兆があった場合に運転員に対してアラートを出力する。運転員はこの出力内容とプロセスデータ等から、プラントの停止などの運転操作を行うべきか判断する。

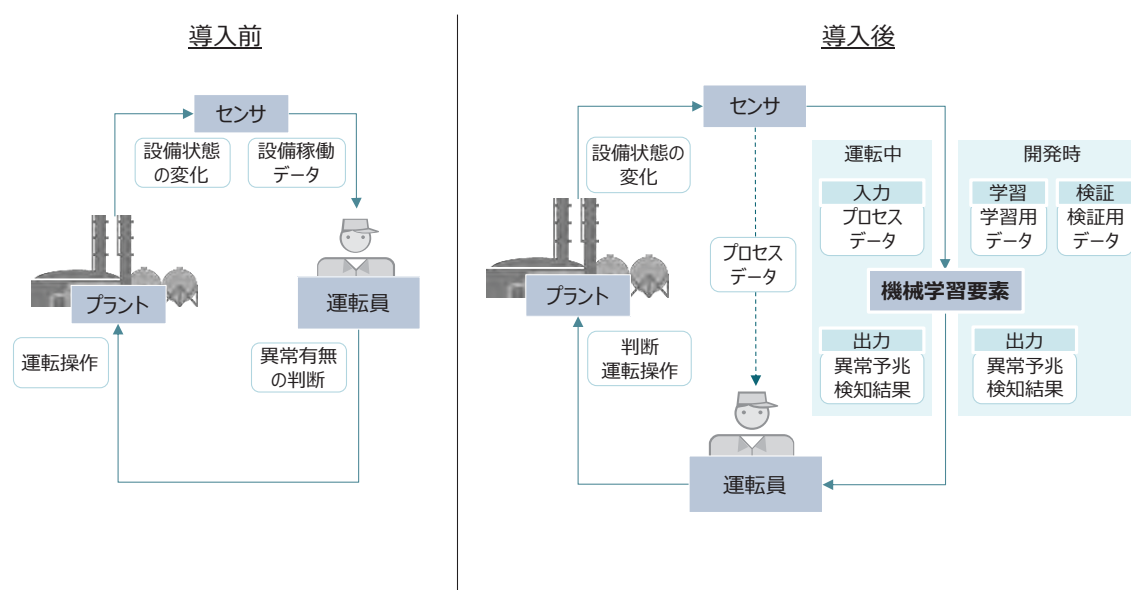


図 3-14 ユースケース「異常予兆検知・診断」の導入イメージ（例）

d. 他のシステムとの関係

本ユースケースでは、機械学習要素と他のシステムとの関係を図 3-15 のように設定している。機械学習要素は、プロセスデータを入力として異常予兆有無を判定し、異常予兆有りと判断された場合に運転員に対してアラートを提示する。運転員は機械学習要素の判断を参考に、各種センサ等の従来監視しているデータを含めて確認し、プラントの停止を含む運転状態の変更の判断を行う。機械学習要素の判断（近未来の異常の有無）そのものを監視・補正外部安全機構は存在せず、別途、プラントの安全性を担保する機械学習利用システムと独立の安全関連系（インターロック等）が存在していることを想定する。

このため、本ユースケースの前提の下では、「AISL 表」を参照する際に「②」の列を適用することとなる。

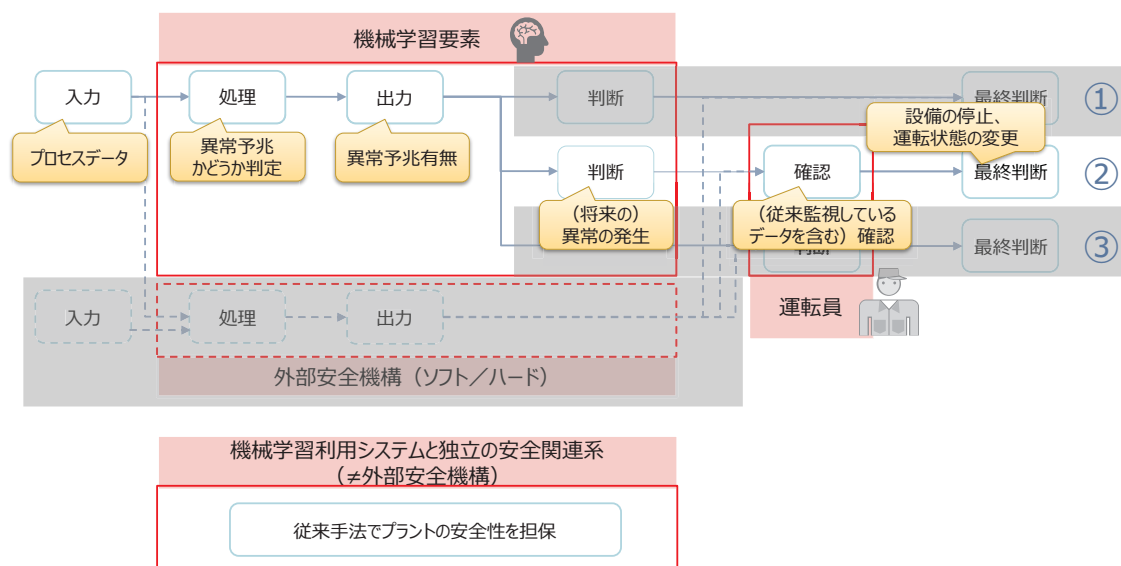


図 3-15 ユースケース「異常予兆検知・診断」における機械学習要素と他のシステムの関係（例）

e. 機械学習要素の構成

本ユースケースでは、機械学習要素の構成を表 3-11 のように設定している。プラントでは異常の発生頻度が低く、多量・多種の異常時のデータの収集が難しいため、学習モデルとして、正常時のプラントデータを用いて正常領域を学習する教師なしの分類モデルを想定する。したがって、正常領域に属さない入力を異常予兆として検出する。学習データとしては正常時のプロセスデータの実測値を用い、テストデータには正常時のプロセスデータ（実測値）と異常時のプロセスデータを用いると想定する。異常時のデータは実測値のみでは十分なテストができないため、実測値に加えて、シミュレーションで生成したデータを用いると想定する。

表 3-11 「異常予兆検知・診断」の機械学習要素の構成（例）

学習手法	分類(教師なし)
学習モデル	正常領域を学習
運用時の入力データ	設備のプロセスデータ
開発時の学習データ	正常時の設備のプロセスデータ(実測値)
開発時のテストデータ	正常時の設備のプロセスデータ(実測値) 異常時の設備のプロセスデータ(実測値+シミュレーションで作成)

2) 利用時品質・外部品質の項目例

本ユースケースの設定を前提とし、利用時品質・外部品質の項目を表 3-12 の通り設定する。2 章に示した階層的な品質確保の手順（図 3-16）においては、「①機械学習を用いたシステムで実現したいことを設定」「②機械学習要素に求める出力を設定し、その達成レベルを決める」に該当する。

表 3-12 ユースケース「異常予兆検知・診断」の利用時品質・外部品質の項目例

利用時品質	外部品質
リスク回避性	
様々なプラントの条件下で、将来の異常の発生を正しく検知する(S-U1)	「異常予兆あり」の場合に、「正常」であると判定する誤判定率を限りなく小さくする(S-E1)
様々なプラントの条件下で、異常発生箇所・異常の深刻度・異常と関連のある変数を正しく出力する(S-U2)	「異常予兆あり」の場合に、異常発生箇所・異常の深刻度・異常と関連のある変数を誤って認識する誤認識率を限りなく小さくする(S-E2)
パフォーマンス	
発報内容の確認に運転員や点検要員の時間を多く割かなくてよい適度な発報頻度とする(P-U1)	誤検知の頻度を一定以下にする(P-E1)
発報を受けてから事故回避の対応が可能な程度に十分に早いタイミングに発報する(P-U2)	決められた時間までに検出する(P-E2)

注) 各項目の記号は、項目間の関係を明らかにするために説明の都合上付与しているものであり、本ガイドラインとして付与を求めるものではなく、レベルの大小とも無関係である。なお、S: Safety、P: Performance、U: Use、E: External を意味する。

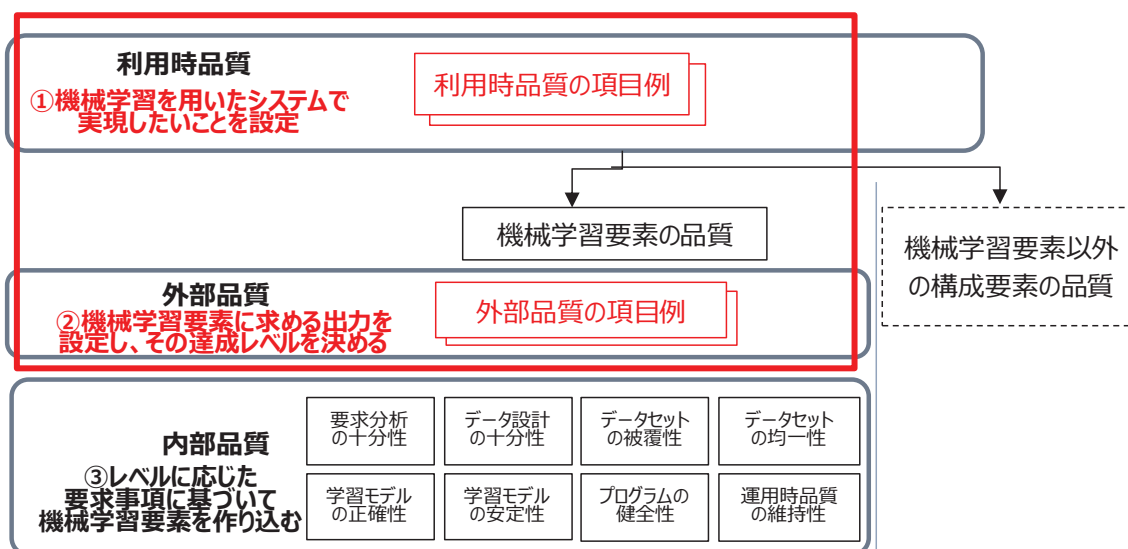


図 3-16 階層的な品質確保における利用時品質・外部品質項目設定の位置づけ

a. 「リスク回避性」に関する検討例

機能要件「異常発生時に発報する」「異常箇所、異常の深刻度、異常と関連のある変数を出力する」を踏まえ、「リスク回避性」の観点からは、機能要件が達成されないことで人的・経済的被害が生じることを防ぐための品質管理を行うために、管理対象である利用時品質及び外部品質を具体的に設定する。

● 利用時品質

人的・経済的被害の予防という観点からは、異常の発生を見逃すことや、異常の内容(箇

所等)を誤って出力することは望ましくない。これらのリスクを回避するために、「リスク回避性」属性の利用時品質として「様々なプラントの条件下で、正しく異常の発生を検知する(S-U1)」と「様々なプラントの条件下で、異常発生箇所・異常の深刻度・異常と関連のある変数を正しく出力する(S-U2)」の2項目を設定する。

● 外部品質

それぞれの利用時品質に対し、機械学習要素のアウトプットに求める外部品質を定める。「様々なプラントの条件下で、異常の発生を正しく検知する(S-U1)」という利用時品質に対応する外部品質は、「『異常予兆あり』の場合に、『正常』であると判定する誤判定率を限りなく小さくする(S-E1)」と設定する。ここで「誤判定率を一定以下とする」ではなく「誤判定率を限りなく小さくする」としているのは、テストで利用できる過去の異常データの量・種類が限られていることを踏まえると、少なくとも、過去の限られた異常データの実績値及び明らかに異常である状況をシミュレートしたテストデータに対しては、正しく異常と判定されることを期待するためである。

また、「様々なプラントの条件下で、異常発生箇所・異常の深刻度・異常と関連のある変数を正しく出力する(S-U2)」という利用時品質に対応する外部品質は「『異常予兆あり』の場合に、異常発生箇所・異常の深刻度・異常と関連のある変数を誤って認識する誤認識率を限りなく小さくする(S-E2)」と設定する。運転員が認識する「関連」を、機械学習要素のアウトプットに置き換え、「関連」と表現している。

それぞれの外部品質に対し、必要な外部品質のレベル「AISL」を設定する。「異常発生時に、『正常』であると判定する誤判定率を限りなく小さくする(S-E1)」については、異常を見逃した場合に想定される人的・経済的被害の大きさを検討し、基準に従って AISL を設定する。同様に、「『異常予兆あり』の場合に、異常発生箇所・異常の深刻度・異常と関連のある変数を誤って認識する誤認識率を限りなく小さくする(S-E2)」については、誤った出力があった場合に想定される人的・経済的被害の大きさを検討し、基準に従って AISL を設定する。

これらのうち最も大きい AISL が、機械学習要素の AISL となり、これに応じて内部品質の要求レベルが決まる。

b. 「パフォーマンス」に関する検討例

機能要件「異常発生時に発報する」「異常箇所、異常の深刻度、異常と関連のある変数を出力する」を踏まえ、「パフォーマンス」の観点からは、機能要件を望ましい水準で達成するための品質管理を行うために、管理対象である利用時品質及び外部品質を具体的に設定する。

● 利用時品質

異常予兆検知・診断では、発報の精度やタイミングを望ましい水準とすることが求められる。発報の精度については、一定の誤発報(正常を異常と誤認識する)は許容しつつも、その頻度が多すぎると確認作業やプラント停止時間が増え、プラントのオペレーションへの悪影響が大きくなるため、許容できない。そのため、利用時品質としては「発報内容

の確認に運転員や点検要員の時間を多く割かなくてよい適度な発報頻度とする(P-U1)」という設定となる。発報のタイミングについては、実際の異常発生直前（数秒前等）に発報しても対応が困難であるため、予兆の検知として意味のある十分に早いタイミングでの発報が求められる。すなわち、利用時品質としては「発報を受けてから事故回避の対応が可能な程度に十分に早いタイミングに発報する(P-U2)」という設定となる。

● 外部品質

それぞれの利用時品質に対し、機械学習要素のアウトプットに求める外部品質を定める。「発報内容の確認に運転員や点検要員の時間を多く割かなくてよい適度な発報頻度とする(P-U1)」という利用時品質に対応する外部品質は、「誤検知の頻度を一定以下にする(P-E1)」と設定する。運転員が認識する「適度な発報頻度」を、機械学習要素のアウトプットに置き換え、「誤検知の頻度を一定以下にする」と表現している。

また、「発報を受けてから事故回避の対応が可能な程度に十分に早いタイミングに発報する(P-U2)」という利用時品質に対応する外部品質は、「決められた時間までに検出する(P-E2)」と設定する。

それぞれの外部品質に対し、必要な外部品質のレベル「AIPL」を設定する。「誤検知の頻度を一定以下にする(P-E1)」については、求める正答率の水準や、必須とする度合いを検討し、基準に従って AIPL を設定する。同様に、「決められた時間までに検出する(P-E2)」については、求める出力タイミングの水準や、必須とする度合いを検討し、基準に従って AIPL を設定する。これらのうち最も大きい AIPL が、機械学習要素の AIPL となり、これに応じて内部品質の要求レベルが決まる。

3) 内部品質確保のための「ユースケース固有の観点」

本ユースケースの設定を前提とした場合の、内部品質の各要求事項を実現するための留意点（「観点」）を表 3-13 に示す。以下に記載する「観点」は、本ユースケースに近い機械学習要素を開発する場合に、参考にすることができる。2 章に示した階層的な品質確保の手順（図 3-17）においては、「③レベルに応じた要求事項に基づいて機械学習要素を作り込む」に該当する。

表 3-13 「異常予兆検知・診断」に係る「ユースケース固有の観点」

※本ケースの内部品質の要求事項と観点の一覧は附録のチェックリストを参照

内部品質	要求事項 ⁵⁹	ユースケース固有の観点
要求分析の 十分性	(要求事項共通)	・ どの箇所のどのような異常を対象とするのかによって、「データ設計の十分性」「データセットの被覆性」の評価にも影響するため、どの箇所のどのような異常を対象とするかという点まで要求を特定する。 ・ 異常の検知と、それに関連する変数との間で、工学的に意味のある因果関係が不明であっても、相関のみで活用することは妨げない。⁶⁰
データ設計 の十分性	・ (Lv1)また、特に重要と考えられる環境要因の差異に対する属性を抽出し、大きなリスクの要因との組み合わせに対応するケースを用意すること。	・ 「環境要因」とは、ここでは異常の検知に影響する環境要因(生産負荷、生産ロット等)を指す。
	(要求事項共通)	・ データセットをシミュレーションで取得する場合は、シミュレータの妥当性を十分検証する。
データセット の被覆性	・ (Lv1)テスト用データセットの取得源や方法を検討し、応用の状況に対して偏りが無いことを期待できるようにすること。	・ 「応用の状況」とは、ここでは検知したい異常の深刻度(ハイアラーム/ハイハイアラーム 等)や、機械学習利用システムの活用場面(常時/一時、昼間/夜間、定常時/非定常時等)を指す。
	・ (Lv1)各ケース毎に、元データから偏りのないサンプル抽出などを行い、偏りが無いことを期待できるようにすること。	・ 本ケースの場合、学習データとして異常時データを網羅することは必須としない。一方で、正常領域における網羅的なサンプル抽出が必要。
	(要求事項共通)	・ 正常時データが実際に正常時のものであることの確認は、その判断が適切に行える専門性を持った要員によって行う。
データセット の均一性	(要求事項共通)	・ 正常データとして想定する様々な範囲(昼間/夜間、定常時/非定常時、季節の違い等)のデータを偏りなく取得する。 ・ ある範囲の正常データが十分取得できない場合、その範囲では異常を検出する精度が低下する可能性があることを認識する。

⁵⁹ 表中では、「機械学習品質マネジメントガイドライン 第1版」記載の要求事項について、「ユースケース固有の観点」に関連するもののみ抜粋している。本表に記載していない項目も、要求事項に含まれる。

⁶⁰ 異常予兆検知・診断においては、「異常の原因の特定と、その原因と異常の因果関係の工学的説明がないと活用できない」、との声がある。一方で、因果関係が未知の相関を見出すことが機械学習の有用性であるという側面もある。本ガイドラインでは、機械学習の活用促進の観点から、ユーザー・ベンダー間の合意のもと、安全を確保する前提で、工学的に意味のある因果関係が不明であっても、相関のみで活用することを妨げないという立場をとっている。

内部品質	要求事項 ⁵⁹	ユースケース固有の観点
機械学習モデルの正確性	(Lv1)テスト段階において一定量の誤判断を許容する場合 (false negative / false positive で扱いを変える場合を含む)については、その判定基準を合理的に事前に決定し、記録しておくこと。	本ケースの場合、誤検知は一定量許容されうるが、テストに使用できる異常データの量・種類に限られるため、見逃しの発生率は極限まで0に近づけることが望ましい。
機械学習モデルの安定性	—	—
プログラムの健全性	—	—
運用時品質の維持性	(要求事項共通)	- 化学系プラントでは特に外部環境の変化(日照条件、風向きなど)の機械学習要素への影響が大きいため、隣接する設備の撤去や変更など、対象設備の直接的変更でない場合でも、対象設備の外部環境に影響する変更に留意する。 - 対象設備の生産負荷に応じた経年劣化の進行を想定し、精度検証、学習モデルのチューニングの頻度を設計する。 (経年劣化ではなく)対象設備を大規模に修繕した場合は、その都度精度検証、学習モデルのチューニングが必要とされる。

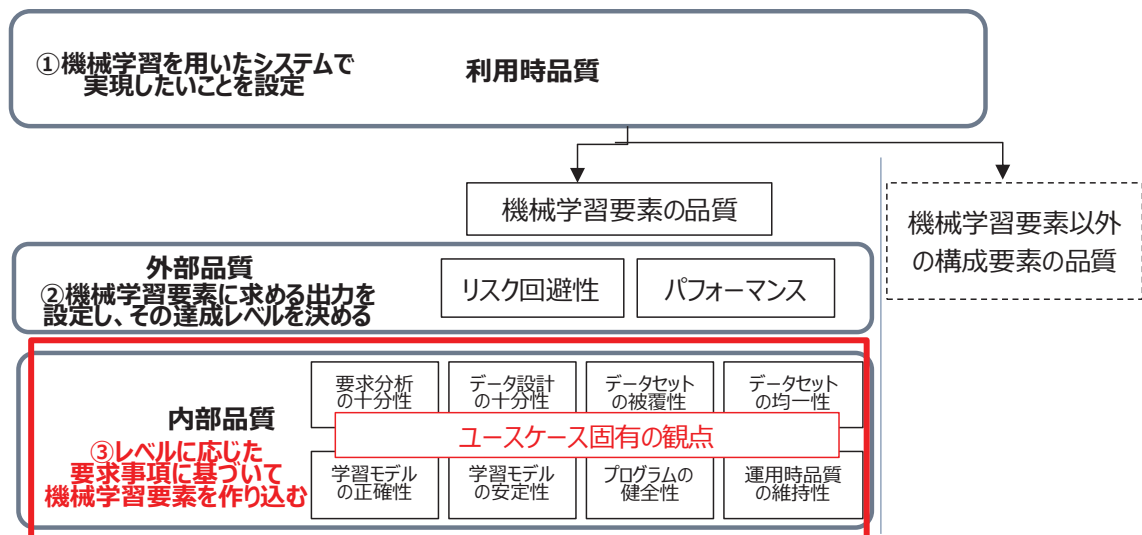
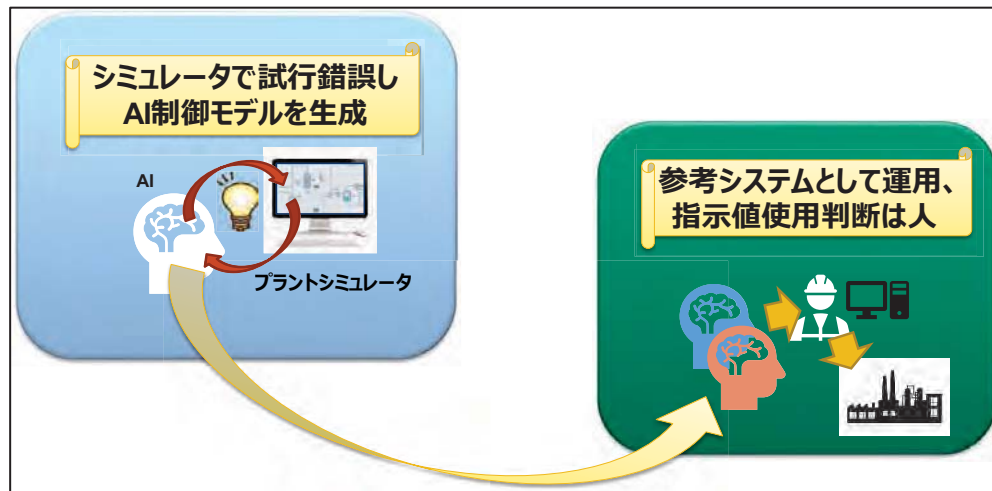


図 3-17 階層的な品質確保における「ユースケース固有の観点」の位置づけ

3.3.5 運転最適化

※本事例は、「運転最適化」のイメージとして掲載するものであり、以降のユースケースの記述の前提ではない。

関連事例(定常運転)：ブタジエン生産プラントにおける AI 制御システム構築
(横河電機株式会社)

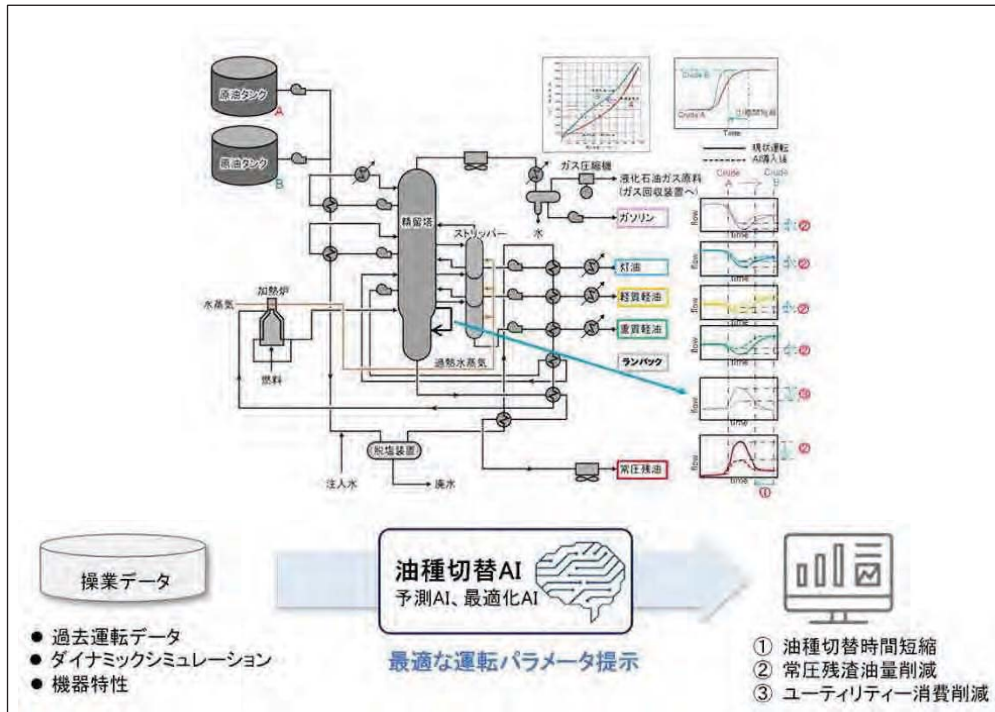


ブタジエン精製プラントでは PID 制御や多変数モデル予測制御により、プラント運転の自動化・安定化を行っている。しかし、天候変化等の外乱等の影響が大きい一部の工程では完全な自動制御は難しく、オペレータによる手動操作も含めてプラントをコントロールしているのが現状である。

そこで、当該プラントのシミュレータを構築して強化学習型 AI 制御アルゴリズムの自動学習を行い、そこで得たモデルを実プラントに適用することで、対象工程の自動化を既存の制御手法では難しい範囲にまで拡大し、人的ミスを減少させるとともに、生産性や運転精度を向上する。

※本事例は、「運転最適化」のイメージとして掲載するものであり、以降のユースケースの記述の前提ではない。

関連事例(非定常運転): 製油所における原料原油切り替え運転最適化 AI
(千代田化工建設株式会社)⁶¹



石油精製プラントでは、輸入原油のタンクが空に近づくたびに、次の油種との性状の違いに応じてベテラン運転員が連続的に十以上のパラメータを同時に調整する切り替え運転を行っている。この運転は3日に一度程度、半日程度の手動運転が必要で、頻度と難易度が高く、操作を誤ると設備の損傷や運転のアップセットに繋がりがうる。したがって、仮に最適運転ができれば生産性と保安力の向上効果が非常に大きい。

動的プラントシミュレータを用いて深層強化学習を行うことで、AIが最適な運転パラメータを出力できる。油種切り替え作業中にリアルタイムで最適な運転パラメータを運転員に提示し、省エネ・製品ロス最小化・早期切り替え完了・安全な運転を実現する。

深層強化学習では、熟練運転員の知見や過去の運転データなどから構築した、運転状況に対する最適化目標(省エネ・製品ロス最小化など)の評価を活用する。動的プラントシミュレータとAIが連携して、様々な運転状況に対して評価が高くなる運転パラメータを学習することで、新しい条件下でも連続的により良い運転パラメータを見つけられるようになる。

⁶¹ 本事例は、「プラントにおける先進的 AI 事例集 ～AI プロジェクトの成果実現と課題突破の実践例～」に詳細を記載している。併せて参照されたい。

1) ユースケースの前提

a. 概要

ユースケース「運転最適化」は、製品の生産性最大化・製品ロス最小化・省エネなどの最適化目標の実現を目的に、プラントの最適化目標に応じた最適な運転パラメータを提示する機械学習利用システムである。プラントの運転は大きく定常運転と非定常運転に分けられる。プラントのスタートアップ・シャットダウン、製造製品切り替え、突発停止時などが非定常運転に該当し、それ以外の比較的安定した状態での運転が定常運転となる。本ユースケースは、定常運転と非定常運転の双方に共通する、プラントの各設備の稼働状況、生産データ、天気・原料・各種条件などから、運転員に対して最適な操作パラメータを提示する機能を想定する。定常運転・非定常運転のいずれかのみに当てはまる内容については、定常運転・非定常運転の別を明示する。

b. 機能要件

本ユースケースでは、機械学習利用システムの機能要件として、「目的に応じた最適な操作パラメータを提示する」を設定する。最適化の目的は、定常運転の場合は生産性向上、非定常運転の場合は定常運転状態への早期移行等を想定する。

c. 導入イメージ

本ユースケースでは、導入のイメージを図 3-18 のように設定している。従来のプラントの運転では、運転員がプラントに設置された各種センサのデータから設備や環境の状態を把握し、過去の経験に基づいて運転目標に応じて運転パラメータの調整を行っていた。一方、機械学習利用システムの導入後は、機械学習要素がプラントの設備や環境のデータからリアルタイムで最適な運転パラメータを算出し、運転員に提示する。運転員は提示された運転パラメータを参考に、センサデータと自身の経験から、プラントに対する運転操作を行う。

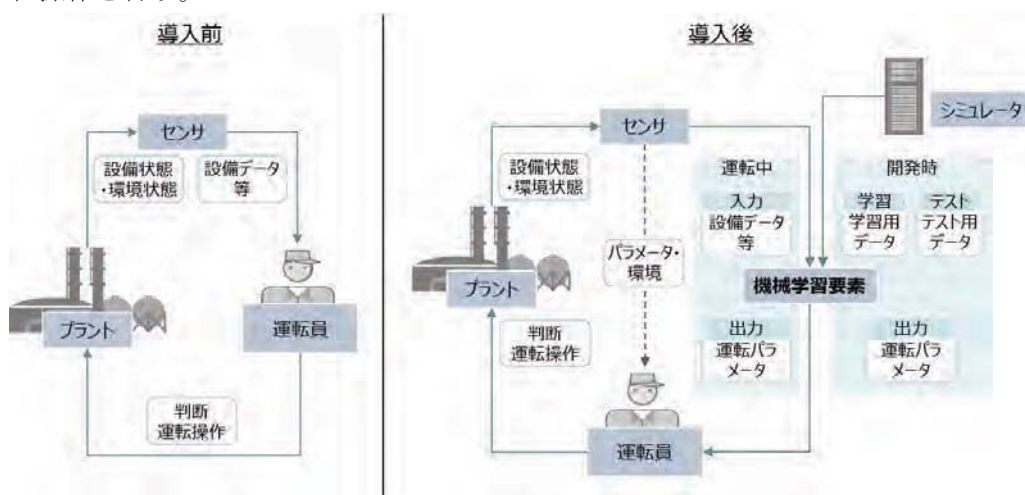


図 3-18 ユースケース「運転最適化」の導入イメージ（例）

d. 他のシステムとの関係

本ユースケースでは、機械学習要素と他のシステムとの関係を図 3-19 のように設定している。機械学習要素は、プラントの各種センサデータ（設備データ、環境データ、生産データなど）を入力とする。設備データは設備の回転数や温度など、環境データは気温や湿度など、生産データは生産量などを想定する。それらの入力から最適化目標に応じて運転パラメータの最適値を算出し、運転員に対して出力する。運転員は自身の経験及び各種センサデータを参照し、機械学習要素が出力した運転パラメータが適当かどうか判断を行い、実際にプラントに対して操作する運転パラメータを決定する。このため、本ユースケースの前提の下では、「AISL 表」を参照する際に「③」の列を適用することとなる。

機械学習要素の出力（運転パラメータの最適値）そのものを監視・補正する外部安全機構として、現在の設備状態が危険水準に到達しているかを判定し、機械学習要素が出力する運転パラメータを監視する処理が存在することを想定する。また、プラントに従来から存在する警報システムや緊急停止システムなどの既存システムも独立して存在していることを想定する。

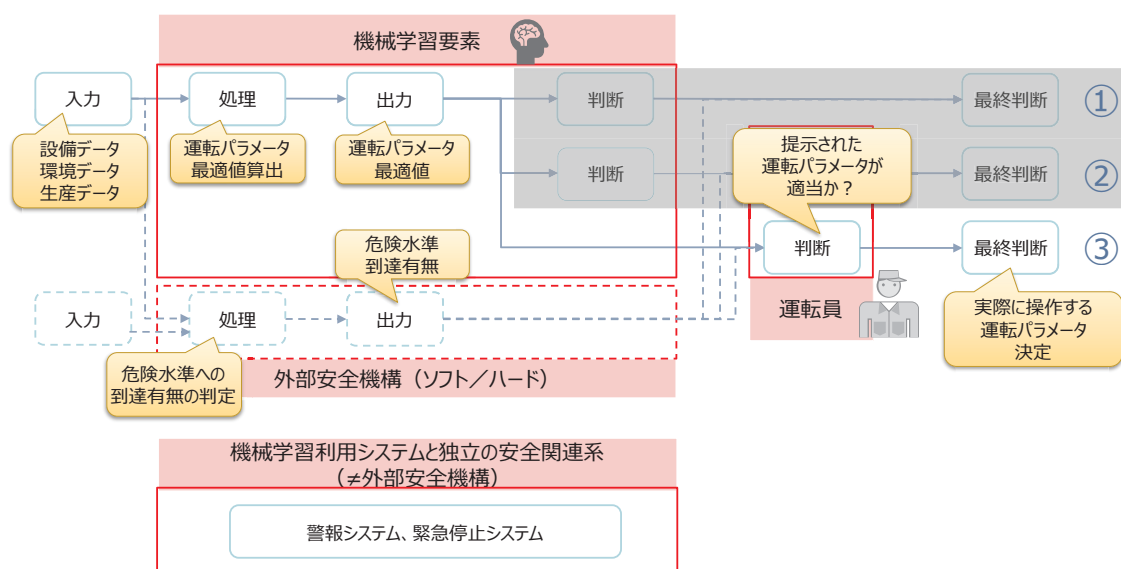


図 3-19 ユースケース「運転最適化」における
機械学習要素と他のシステムの関係（例）

e. 機械学習要素の構成

本ユースケースでは、機械学習要素の構成を表 3-14 のように設定している。学習モデルとして、最適な運転パラメータを学習する強化学習モデルを想定する。強化学習では、設備のパラメータ上限付近での運転など、実際のプラントでの実測値のみで学習データ・テストデータの被覆性と十分性を満たすことは難しいため、シミュレータの活用を前提とする。

表 3-14 「運転最適化」の機械学習要素の構成（例）

学習手法	強化学習
学習モデル	最適な運転パラメータを学習する強化学習
運用時の入力データ	設備データ・環境データ・生産データ
開発時の学習データ	設備データ・環境データ・生産データ+運転パラメータ(実測値+シミュレーションで作成)
開発時のテストデータ	設備データ・環境データ・生産データ+運転パラメータ(実測値+シミュレーションで作成)

2) 利用時品質・外部品質の項目例

本ユースケースの設定を前提とし、利用時品質・外部品質の項目を表 3-15 の通り設定する。2 章に示した階層的な品質確保の手順（図 3-20）においては、「①機械学習を用いたシステムで実現したいことを設定」「②機械学習要素に求める出力を設定し、その達成レベルを決める」に該当する。

表 3-15 ユースケース「運転最適化」の利用時品質・外部品質の項目例

利用時品質	外部品質
リスク回避性	
設備の安全上の許容動作スペックを超えるような運転状態をもたらさない(S-U1)	最適化を行う設備のパラメータ変数の範囲を、安全上の許容動作スペックに対応する範囲に制限する(S-E1)
パフォーマンス	
(定常運転)生産性を向上させるパラメータを提示する(P-U1)	(定常運転)生産量を一定割合向上させるパラメータを提示する(P-E1)
(非定常運転)定常運転状態に早期移行させる等のパラメータを提示する(P-U1-2)	(非定常運転)定常運転状態に移行する時間を一定割合短縮させる等のパラメータを提示する(P-E1-2)

注) 各項目の記号は、項目間の関係を明らかにするために説明の都合上付与しているものであり、本ガイドラインとして付与を求めるものではなく、レベルの大小とも無関係である。なお、S: Safety、P: Performance、U: Use、E: External を意味する。

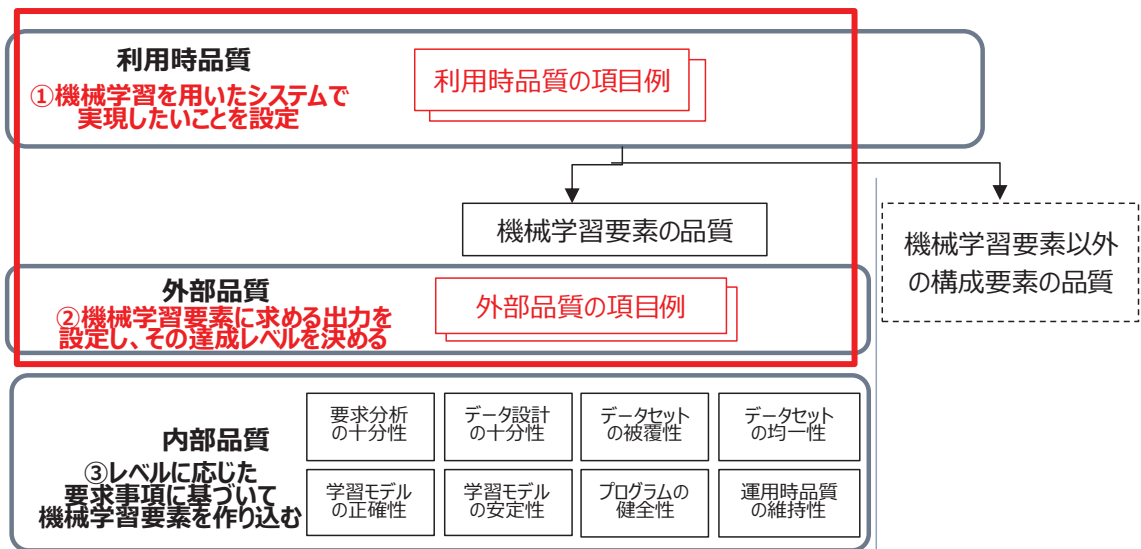


図 3-20 階層的な品質確保における利用時品質・外部品質項目設定の位置づけ

a. 「リスク回避性」に関する検討例

機能要件「目的に応じた最適な操作パラメータを提示する」を踏まえ、「リスク回避性」の観点からは、機能要件が達成されないことで人的・経済的被害が生じることを防ぐための品質管理を行うために、管理対象である利用時品質及び外部品質を具体的に設定する。

● 利用時品質

人的・経済的被害の予防という観点からは、出力する運転パラメータが設備の想定を超えるものになることは望ましくない。このリスクを回避するために、「リスク回避性」属性の利用時品質として「設備の安全上の許容動作スペックを超えるような運転状態をもたらしさない(S-U1)」という項目を設定する。

● 外部品質

設定した利用時品質に対し、機械学習要素のアウトプットに求める外部品質を求める。「設備の安全上の許容動作スペックを超えるような運転状態をもたらしさない(S-U1)」という利用時品質に対応する外部品質は、「最適化を行う設備のパラメータ変数の範囲を、安全上の許容動作スペックに対応する範囲に制限する(S-E1)」と設定する。

この外部品質に対し、必要な外部品質のレベル「AISL」を設定する。設備の想定スペックを超える運転パラメータが提示された場合に想定される影響について、外部安全機構や機械学習利用システムと独立の安全関連系を考慮して検討し、基準に従って AISL を設定する。

本ユースケースでは、信頼性が確保された外部安全機構および機械学習利用システムと独立の安全関連系（緊急停止システム等）が安全を確保することを前提としているため機械学習要素には高度な安全性を求める必要が無いことを考慮し、SIL 評価を行って「SIL なし」とした上で、AISL 表に基づいて、AISL0～0.2 程度が割り当てられると考える。これは、あくまで本ガイドラインで例示した前提のもとで判断されるものであり、本ユース

ケースに類似するあらゆるケースで同様の評価となることを意味しない。機械学習要素に求める安全機能の水準は、個別の適用ごとに判断する必要がある。

b. 「パフォーマンス」に関する検討例

機能要件「目的に応じた最適な操作パラメータを提示する」を踏まえ、「パフォーマンス」の観点からは、機能要件を望ましい水準で達成するための品質管理を行うために、管理対象である利用時品質及び外部品質を具体的に設定する。

● 利用時品質

運転最適化では、最適化目標である生産性の向上（定常運転）、定常運転状態への早期移行等（非定常運転）を達成することが求められる。そのため、利用時品質としては「生産性を向上させるパラメータを提示する(P-U1-1)」「定常運転状態に早期移行させる等のパラメータを提示する(P-U1-2)」と設定する。

● 外部品質

設定した利用時品質に対し、機械学習要素のアウトプットに求める外部品質を定める。定常運転の場合、「生産性を向上させるパラメータを提示する(P-U1-1)」という利用時品質に対応する外部品質は、「生産量を一定割合向上させるパラメータを提示する(P-E1)」と設定する。非定常運転の場合、「定常運転状態に早期移行させるパラメータを提示する(P-U1-2)」という利用時品質に対応する外部品質は、「定常運転状態に移行する時間を一定割合短縮させるパラメータを提示する(P-E1-2)」と設定する。

3) 内部品質確保のための「ユースケース固有の観点」

本ユースケースの設定を前提とした場合の、内部品質の各要求事項を実現するための留意点（「観点」）を表 3-16 に示す。以下に記載する「観点」は、本ユースケースに近い機械学習要素を開発する場合に、参考にすることができる。2 章に示した階層的な品質確保の手順（図 3-21）においては、「③レベルに応じた要求事項に基づいて機械学習要素を作り込む」に該当する。

表 3-16 「運転最適化」に係る「ユースケース固有の観点」

※本ケースの内部品質の要求事項と観点の一覧は附録のチェックリストを参照

内部品質	要求事項 ⁶²	ユースケース固有の観点	
		定常運転	非定常運転
要求分析の 充分性	—	—	

⁶² 表中では、「機械学習品質マネジメントガイドライン 第1版」記載の要求事項について、「ユースケース固有の観点」に関連するもののみ抜粋している。本表に記載していない項目も、要求事項に含まれる。

内部品質	要求事項 ⁶²	ユースケース固有の観点	
		定常運転	非定常運転
データ設計の 十分性	(Lv1)また、特に重要と考えられる環境要因の差異に対する属性を抽出し、大きなリスクの要因との組み合わせに対応するケースを用意すること。	「環境要因」とは、ここでは運転手順や原材料などを指す。	
	(要求事項共通)	データセットをシミュレーションで取得する場合は、シミュレータの妥当性を十分検証する。	
データセット の被覆性	(Lv1)テスト用データセットの取得源や方法を検討し、応用の状況に対して偏りが無いことを期待できるようにすること。	「応用の状況」とは、ここでは適用する操作場面(季節、時間帯等)、操作対象設備などを指す。	「応用の状況」とは、ここでは適用する操作場面(季節、時間帯、スタートアップ/シャットダウン等)、操作対象設備などを指す。
	(要求事項共通)	- 天候などの「外乱」を想定したデータセットになっているか留意する。 - ベテランオペレータの運転を学習する際は、偏ったケース設定となっていないか留意する。	
データセット の均一性	—	—	
機械学習モデルの 正確性	テスト段階において一定量の誤判断を許容する場合(false negative/false positive で扱いを変える場合を含む)については、その判定基準を合理的に事前に決定し、記録しておくこと。	「機械学習要素による操作パラメータ提示に安全機能を期待しない(設備の想定スペックを超えるような操作パラメータ提示に対しては、外部安全機構及びオペレータ判断により操作に反映しない)前提」の場合は、「危険な操作につながるパラメータの出力の許容水準を極限まで0に近づける」といった検討は求めない。SIL 評価等に基づき、機械学習要素と他のシステムの安全機能の分担を適切に考慮して機械学習要素への要求水準を決定する。	
	(要求事項共通)	強化学習を用いたとしても、運用開始前にテストを実施して「機械学習モデルの正確性」の要求事項を満たす。⁶³	
機械学習モデルの 安定性	—	—	

⁶³ 強化学習の場合、テストを実施せずに運用開始し、運用しながらより良い性能を追求することも想定され、この場合は「機械学習モデルの正確性」の要求事項を適用できない。ただし、プラント保安分野では、運用開始時には一定の性能を達成していることをテストで確認することが必要であることから、本ガイドラインでは、強化学習を用いたとしても、運用開始前にテストを実施して「機械学習モデルの正確性」の要求事項を満たすことを求めている。

内部品質	要求事項 ⁶²	ユースケース固有の観点	
		定常運転	非定常運転
プログラムの健全性	—	—	
運用時品質の維持性	(要求事項共通)	- 最適値の評価は定期的／継続的にモニターし、異常がないか確認する。 - 設備や運転手順に変更が行われた場合は機械学習要素の出力への影響があるため、モデルを更新する。 - 最適化の目標に対して極限まで運転条件を追求すると、安定性が損なわれることがあるため、機械学習要素の出力の範囲を限定するなどの対応を行う。 - 想定していた原料(原油種等)の内挿範囲内で運転されていることを確認する。 - 運用中の設備の諸条件(反応初期／終期、運転条件、原料、品質要求、スタートアップ・シャットダウン操作の許容時間等)を考慮して機械学習要素の出力の品質を確認する。	

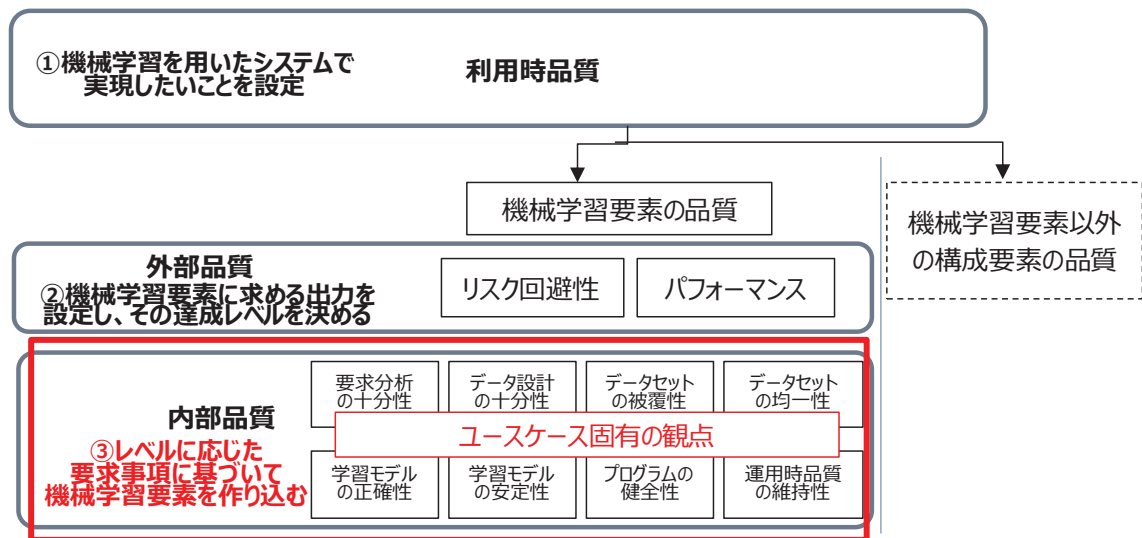


図 3-21 階層的な品質確保における「ユースケース固有の観点」の位置づけ

4. ガイドラインの活用の流れ

ここまで、2 章でプラント保安分野における機械学習の信頼性評価の具体的な方法を説明し、3 章でユースケースに基づく例を紹介した。本章では、信頼性評価の具体的なステップに即して2 章および3 章を活用する方法を示す。

まず、本ガイドラインの活用の主体について、どういった役割・事業者がどういった目的で本ガイドラインを参照できるか整理する（4.1）。またシステム構築・運用において、各主体（保全担当、プラントシステム担当等）が信頼性評価において具体的に実施する事項、実施に際しての本ガイドラインの参照箇所等、信頼性評価におけるガイドラインの適用の具体的な流れを説明する（4.2）。

4.1 ガイドラインの活用の主体

本ガイドラインに基づく信頼性評価は、機械学習利用システムの要求・要件定義～運用の各プロセスで行われる。このプロセスにおいては、複数の企業の多様な主体が関与する。関与する主体の人数や所属企業、役割分担について、プラント保安分野において一般的に必要な役割を表 4-1 に示す。各担当はそれぞれの立場から機械学習利用システムの信頼性確保に関与する。

表 4-1 機械学習利用システムの開発プロジェクトの担当

部署・役割	プロジェクトにおける担当内容
経営企画	プロジェクトの実行に関する意思決定 ※直接のガイドライン活用主体ではない
事業企画担当	機械学習利用システムの開発プロジェクトの主担当(予算、スケジュール等)
品質保証担当 (機械学習利用システム・機械学習要素)	機械学習利用システム全体の利用時品質の評価・確認 機械学習要素の外部品質・内部品質の評価・確認
環境安全担当	機械学習利用システムの安全に関わる内容のレビュー ⁶⁴
現場担当 現場の管理担当 (製造担当、設備管理担当)、 現場の作業担当 (運転員、保全員)	運転・保全に用いるシステムの利用時品質のレビュー 機械学習要素に求める外部品質のレビュー 運用時に所期の効果が出ているかレビュー 外部安全機構や安全関連系のレビュー
プラントシステム担当	機械学習利用システムの利用時品質、機械学習要素の外部品質の設定 機械学習利用システムの機械学習要素以外の構成要素(外部安全機構等)や安全関連系の整理 機械学習要素の開発に係るレビュー・データの提供
機械学習設計・開発担当	機械学習要素の設計・開発

プラント保安分野では、運転・保全を直接担っている担当者(表 4-1「現場担当」)が機械学習利用システム全体及び機械学習要素の品質に対してレビューを行うことや、DCS⁶⁵をはじめとするプラントの既存システムの運用管理を担う担当者(表 4-1「プラントシステム担当」)がデータ提供や機械学習利用システムの要件定義等で関与することが特徴的である。

なお、「品質保証担当」は、機械学習利用システムの開発に関わるその他の人員から適切な独立性の水準を持った適切な組織又は人員によって行われることが必要な場合がある。独立性の水準は、プラントが準拠する既存の基準(機能安全規格等)に従って判断する。

⁶⁴ 「レビュー」とは、本ガイドラインを用いて信頼性評価を行う主担当から依頼を受け、自身の業務所掌・専門性に基づいて主担当による検討事項を確認することを意味する。例えば、機械学習利用システムの機能要件や利用時品質を、プロジェクトを牽引する「事業企画担当」が設定し、それが正しく設定されているか、システムを最終的に利用する「現場担当」が確認を行う場合、主担当は「事業企画担当」で、レビューを行うのは「現場担当」となる。なお、レビューを行う担当者は、必ずしもガイドラインの内容を読んで理解する必要は無く、主担当者からの依頼に応える形で信頼性評価に関与する。詳細に実施する事項は、4.2を参照。

⁶⁵ DCS(Distributed Control System:分散形制御システム)とは、プラントの制御システムである。プラント全体を一つの制御装置で制御するのではなく、構成機器毎に制御装置があり、それがネットワークで接続されることにより、大規模なプラントの操業の制御を実現している。

参考として、それぞれの担当者の所属企業の例を表 4-2 に示す。

主として、機械学習利用システムの導入検討・判断を行うユーザー企業（プラントオーナー等）が中心となって開発プロジェクトが遂行されることが想定される。プラント設備に関連したデータ提供等についてはプラント設備のベンダー企業、機械学習利用システム・機械学習要素の設計・開発についてはシステムのベンダー企業・SIer 企業や AI ベンダー企業が参画する場合がある。

表 4-2 機械学習利用システムの開発プロジェクトにおける役割別の所属企業の例

部署・役割		所属する企業の例
経営企画		プラントオーナー企業
事業企画担当		プラントオーナー企業
品質保証担当	システム全体	システムのベンダー企業、SIer 企業 プラントオーナー企業
	機械学習要素	AI ベンダー企業
環境安全担当		プラントオーナー企業
現場担当	製造担当 設備管理担当	プラントオーナー企業 プラント設備のベンダー企業
	運転員	プラントオーナー企業
	保全員	プラントオーナー企業 メンテナンス企業
プラントシステム担当		システムのベンダー企業、SIer 企業 プラントオーナー企業
機械学習設計・開発担当		AI ベンダー企業

各担当者の役割に応じて、ガイドラインの活用場面が異なるため、3つの段階で設定される品質の中で重点的に確認すべき品質が異なる。各担当の確認すべき品質について、表 4-3 に示す。

それぞれ、特に各品質について具体的な記載のある 2.1 及び 2.2 において、自身が確認すべき品質に関する記載箇所を中心にガイドラインを理解する必要がある。

表 4-3 各担当のガイドラインの活用場面と確認すべき品質

部署・役割	ガイドラインの活用場面	利用時品質	外部品質	内部品質
経営企画	—	○	—	—
事業企画担当	システム全体の目的・利用時品質の設定、経営企画への目的・利用時品質の説明	● →2.1.1 2.1.3 2.2.1	○	○
品質保証担当	利用時品質の確認 外部品質・内部品質の確認	● →2.1.1 2.1.3 2.2.1	● →2.1.2 2.1.3 2.2.2 2.2.3	● →2.1.4 2.2.4 2.2.5
環境安全担当		○	○	○
現場担当	利用時品質に対するレビュー 外部品質に対するレビュー 運用時の効果・品質のレビュー	○	○	○
プラントシステム担当	利用時品質の設定、外部品質の設定、内部品質に係るデータ提供やレビュー	● →2.1.1 2.1.3 2.2.1	● →2.1.2 2.1.3 2.2.2 2.2.3	● →2.1.4 2.2.4 2.2.5
機械学習設計・開発担当	内部品質の把握、内部品質の要求事項に則した機械学習要素の設計・開発・更新	—	○	● →2.1.4 2.2.4 2.2.5

凡例

- : 信頼性評価の手順をガイドラインの記載に則って実施する主体
- : 信頼性評価のレビューに応じる(必ずしもガイドラインの内容を読んで理解する必要は無く、主担当者からの依頼に応える形で信頼性評価に関与する)

4.2 ガイドラインの適用の流れ

機械学習利用システムの構築（要求・要件定義～テスト・検収）及び運用の各フェーズにおいて 4.1 で示した主体が信頼性評価においてどのような役割を果たし、どのようにガイドラインを活用するのかを示す。

4.2.1 担当別・フェーズ別の実施項目

担当・フェーズ別の信頼性評価に係る実施項目の全体像を表 4-4 に示す。

なお、機械学習利用システムが保全か運転かで、参画する主体は異なる。保全に用いる機械学習利用システムでは、現場担当として現場の管理担当である設備管理担当と現場の作業担当である保全員が関わりと想定されるが、運転に用いる機械学習利用システムでは現場の管理担当である製造担当と現場の作業担当である運転員・保全員が関わりと想定される。したがって、本章以降の記述並びに「附録 内部品質確保のための『プラント保安分野での観点』チェックリスト」ではこの点を区別して整理している。

表 4-4 機械学習利用システム構築・運用のフェーズ・担当別の実施項目

経営企画	担当	PoC	要求・要件定義	設計	実装	テスト・検収	運用
事業企画担当			システム開発に関する意思決定			システムの導入に関する意思決定	
			システムの目的の 設定、システムの機能要件・利用時品質のレビュー			検収における品質全体のレビュー	
品質保証	システム品質保証担当					テストを踏まえた外部品質の評価、システムの検収	利用時品質・外部品質の確認
	機械学習品質保証担当				機械学習要素の開発における品質保証		内部品質の確認
環境安全担当							
現場担当				外部品質に係る外部安全機構・安全関連系のレビュー		テストを踏まえた外部品質の評価結果のレビュー	
	現場の管理担当 (保全)設備管理担当 (運転)製造担当		システムの目的・機能要件・利用時品質のレビュー	外部品質及びそのレベル設定のレビュー、外部品質に係る外部安全機構・安全関連系のレビュー	機械学習要素の開発に係るレビュー	検収における品質全体のレビュー	利用時品質の確認結果のレビュー、システムの更新に係るレビュー・データ提供
	現場の作業担当 (保全)保全員 (運転)運転員・保全員		システムの目的・機能要件・利用時品質のレビュー		機械学習要素の開発に係るレビュー	検収における品質全体のレビュー	利用時品質の確認結果のレビュー、システムの更新に係るレビュー・データ提供
プラントシステム担当			システムの機能要件・利用時品質の設定	外部品質の設定、外部品質のレベル設定に係る機械学習要素以外の構成要素(外部安全機構等)や安全関連系の整理、外部品質のレベル設定	機械学習利用システムの機械学習要素以外の構成要素(外部安全機構等)の開発、機械学習要素の開発に係るレビュー・データ提供	テストを踏まえた外部品質の評価結果のレビュー	外部品質の確認結果のレビュー、機械学習要素以外の構成要素(外部安全機構等)の更新
機械学習設計・開発担当					内部品質のレベル設定、機械学習要素の設定・開発		内部品質に係る機械学習要素の更新

凡例 **太字・下線**：該当する品質に係る品質保証活動の主な実施主体、下線のみ：該当する品質に係る品質保証活動をレビュー等で支援する主体

(注) データの収集等の内部品質の作り込みは「実装」フェーズに含めている。

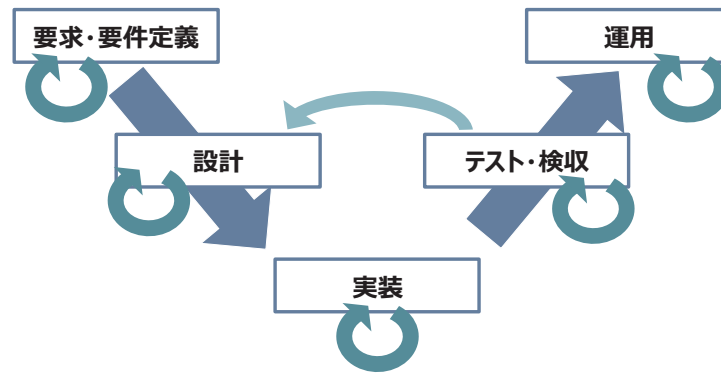
表 4-4 の横軸で示した機械学習利用システムの構築・運用におけるフェーズ別に、信頼性評価に係る実施項目を整理し表 4-5 に示した。

表 4-5 機械学習利用システムの構築・運用におけるフェーズ別の実施項目

フェーズ	ステップ	品質保証での実施項目
PoC	—	—
要求・要件定義	1	システムの目的の設定
	2	システムの機能要件・利用時品質の設定
設計	3	外部品質の設定
	4	外部品質のレベルに係る機械学習利用システムと安全関連系の整理
	5	外部品質のレベル設定
実装	6	内部品質のレベル設定
	7	機械学習要素の設計・開発
	8	機械学習利用システムの機械学習要素以外の構成要素(外部安全機構等)の開発
テスト・検収	9	テスト(外部品質の評価)
	10	検収
運用	11	利用時品質の確認
	12	外部品質の確認
	13	内部品質(運用時品質の維持性)の確認
	14	システムの更新

なお、本ガイドラインではフェーズを時系列で順に行うもの（要求・要件定義→設計→実装→テスト・検収→運用）として表現しているが、本ガイドラインは必ずしもウォーターフォール型の開発スタイルに限って適用できるものではない。図 4-1 に示すように、各フェーズ及びフェーズ間で反復的に行うことも考えられる。例えば、テストフェーズにおいて、設計フェーズで定めた信頼性（例：外部品質として、誤検知頻度を限りなく小さくする）を十分達成していないと評価された場合は、その前のフェーズ・ステップに戻り、なぜ信頼性が十分達成されていないのか順を追って確認・再度実施することが有効である。このように、各フェーズ・ステップの実施事項を適宜行き来しながら信頼性評価を行うことも可能である。

また、本ガイドラインにおいても、機械学習要素の設計・開発（ステップ 7）においては一般にアジャイル型の開発スタイルがとられると想定しており、本ガイドラインの適用により機械学習要素の開発スタイルの変更を要するものではない。



各フェーズ・フェーズ間における反復的な開発

図 4-1 フェーズ間で反復のあるシステム構築・運用のスタイル

4.2.2 実施項目・実施内容の全体像

実施担当者・実施項目及び内容の全体像をリストとして表 4-6 に示す。各フェーズの記載及び以下の表の記載における担当は代表的な例であり、プロジェクトに応じて適切に設置される必要がある。

表 4-6 プラント保安分野における機械学習の信頼性評価 実施項目・内容リスト

(凡例)		実施者：太字…当該実施項目の主な実施主体 細字…当該実施項目のレビューの実施主体)			
フェーズ	ステップ	品質保証での実施項目	実施者	実施内容	ガイドライン
PoC	—	—	—	—	—
要求・要件定義	1	システムの目的の設定	事業企画担当	改善したい「現状の課題」を整理、「システムの目的」を設定し、機械学習利用システムが必要か判断	—
			現場担当	課題やシステムの目的に対し現場・ユーザーの立場からレビュー	
	2	システムの機能要件・利用時品質の設定	プラントシステム担当	機械学習利用システムの機能要件と利用時品質を設定	2.1.1
			事業企画担当	設定された機械学習利用システムの機能要件と利用時品質に対し企画者の立場からレビュー	2.1.3 2.2.1 (3.3)
設計			現場担当	設定された機械学習利用システムの機能要件と利用時品質に対し現場・ユーザーの立場からレビュー	
	3	外部品質の設定	プラントシステム担当	利用時品質を踏まえ、機械学習要素の外部品質を設定	2.1.2
			現場の管理担当	設定された外部品質の具体化内容に対し現場の立場からレビュー	2.1.3 2.2.2 (3.3)
	4	外部品質のレベルに係る機械学習利用システムと安全関連系の整理	プラントシステム担当	機械学習利用システムの各構成要素(機械学習要素、機械学習要素以外の構成要素(外部安全機構等))、同システムと独立した安全関連系の有無や機能要件を整理	2.2.3
			現場の管理担当	機械学習利用システムの機械学習要素以外の構成要素(外部安全機構等)や同システムと独立した安全関連系について現場の立場からレビュー	
			環境安全担当	機械学習利用システムの機械学習要素以外の構成要素(外部安全機構等)や同システムと独立した安全関連系について安全管理の立場からレビュー	

フェーズ	ステップ	品質保証での実施項目	実施者	実施内容	ガイドライン
実装	5	外部品質のレベル設定	プラントシステム担当 現場の管理担当	機械学習要素の外部品質のレベルの設定 外部品質のレベルに対し現場の立場からレビュー	2.2.3
	6	内部品質のレベル設定	機械学習設計・開発担当	外部品質のレベルに応じて、機械学習要素の内部品質のレベルを設定	2.1.4 2.2.4 附録(チェックリスト) (3.3)
	7	機械学習要素の設計・開発	機械学習設計・開発担当 機械学習品質保証担当 現場担当	機械学習要素の設計、内部品質の要求事項と観点に応じた開発 機械学習要素の開発において、設定した内部品質のレベルと観点を満たしているか品質保証の立場から確認 機械学習要素の開発において、現場・ユーザーの立場からデータ提供	2.2.5 附録(チェックリスト)
	8	機械学習要素以外の構成要素(外部安全機構等)の開発	プラントシステム担当 プラントシステム担当	機械学習要素の開発において、データ管理の立場からデータ提供 設定した外部品質・利用時品質に必要な機械学習要素以外の構成要素(外部安全機構等)を開発	—
	9	テスト(外部品質の評価)	システム品質保証担当 プラントシステム担当 環境安全担当	機械学習利用システムのテストを行い、その結果を評価。機械学習要素の外部品質が要求水準に達しているか評価 テスト結果の評価を踏まえて、外部安全機構・安全関連系の整理や、外部品質の設定及びレベル設定をした立場からレビュー テスト結果の評価に対し、安全管理の立場からレビュー	2.2.3
	10	検収	システム品質保証担当 事業企画担当	テストでの外部品質の評価結果から機械学習利用システム全体の評価を行い、それらの評価結果等を踏まえ基準を満たす機械学習利用システムであれば検収 機械学習利用システム全体の評価をプロジェクト発案者の立場からレビュー	—

フェーズ	ステップ	品質保証での実施項目	実施者	実施内容	ガイドライン
運用			現場担当	利用時品質の評価結果に対し現場・ユーザーの立場からレビュー	
	11	利用時品質の確認	システム品質保証担当	運用中の機械学習利用システムの利用時品質を確認	2.1.1
			現場担当	利用時品質の確認結果に対し現場・ユーザーの立場からレビュー	2.1.3 2.2.1
	12	外部品質の確認	システム品質保証担当	運用中の機械学習利用システムの利用時品質の確認結果も踏まえ、外部品質を確認	2.2.3
			プラントシステム担当	外部品質の確認結果を、外部安全機構・安全関連系の整理や、外部品質の設定及びレベル設定をした立場からレビュー	
	13	内部品質(運用時品質の維持性)の確認	機械学習品質保証担当	運用中の機械学習利用システムの利用時品質・外部品質の確認結果も踏まえ、内部品質の要求事項と観点の対応状況を確認	2.2.4 2.2.5 附録(チェックリスト)
14	システムの更新		機械学習設計・開発担当	外部品質・内部品質の確認結果に応じて機械学習要素を更新	2.2.4
			プラントシステム担当	利用時品質・外部品質の確認結果に応じて機械学習要素以外の構成要素(外部安全機構等)を更新、システム管理の立場から機械学習要素の更新に対しデータ提供	2.2.5 附録(チェックリスト)
			現場担当	機械学習要素の更新において、現場・ユーザーの立場からレビュー	

4.2.3 フェーズ別の実施項目・実施内容

(1) PoC

PoC は、純粋な試行的検討から本格開発の準備まで多様な意味を含む概念であり、すべての場合に信頼性評価の検討が求められるものではない。しかし PoC 後半で本格開発を見据える段階においては、以降の開発の際に信頼性評価を円滑に行う準備として、各フェーズにおける信頼性評価の実施事項・留意事項を予め考慮することが有効である。具体的には、表 4-7 に示した事項は、構築プロセスを円滑に進める上で、PoC 終了段階で目処をつけておくことが望ましい。具体的なシステム・設備が未確定な PoC 段階で、外部品質の厳密なレベル設定を行うことは困難だが、機能要件・利用時品質に照らして外部品質の要求レベルの見通しをつけ、内部品質の作り込みの方針を検討しておくことが有効である。

具体的な実施内容は、該当するフェーズの記載を参照されたい。

本格開発においては、それぞれ役割をもった担当が各実施項目を行うが、PoC フェーズにおいては、それらの役割を開発のコアメンバー（事業企画担当、プラントシステム担当、機械学習設計・開発担当等）で仮定しながら検討することが想定される。

表 4-7 PoC 終了段階で目処をつけることが望ましい項目

フェーズ	ステップ	品質保証での実施項目	ガイドライン該当箇所
要求・要件定義 (4.2.3(2))	1 (4.2.3(2)1))	システムの目的の設定	—
	2((2)2))	システムの機能要件・ 利用時品質の設定	2.1.1/2.1.3/2.2.1 (3.3)
設計 (4.2.3(3))	3 (3.2.1(3)3))	<u>外部品質</u> の設定	2.1.2/2.1.3/2.2.2 (3.3)
	5((3)5))	<u>外部品質</u> のレベル設定	2.2.3
実装 (4.2.3(4))	6 (4.2.3(4)6))	<u>内部品質</u> のレベル設定	2.1.4/2.2.4/附録(チェックリスト) (3.3)

注) フェーズ・ステップの () 内の数字は、本ガイドライン 4.2 節内で該当するフェーズ・ステップが記載されている項目番号である。

(2) 要求・要件定義

要求・要件定義においては、まずプラントオーナー企業が自社の課題を特定し、それを解決するシステムの機能要件や利用時品質を設定する。

機械学習の利用の是非は、その課題解決への必要性に応じて判断される。課題を特定した結果、機械学習要素が必要である場合については、本ガイドラインに基づいて信頼性評価を行う。機械学習要素が必要でない場合は、機械学習要素を含まないシステムとして従来通りの信頼性評価を行うこととなる。

表 4-8 フェーズ「要求・要件定義」の実施事項

ステップ	品質保証での実施項目	実施者	ガイドライン該当箇所
1	システムの目的の設定	事業企画担当: システム導入によって改善したい現状の課題を整理、システムの目的を設定 (現場担当: 課題やシステムの目的に対し現場・ユーザーの立場からレビュー)	ー
2	システムの機能要件・利用時品質の設定	プラントシステム担当: 機械学習利用システムの機能要件と利用時品質を設定 (事業企画担当: 設定された機械学習利用システムの機能要件と利用時品質に対し企画者の立場からレビュー) 現場担当: 設定された機械学習利用システムの機能要件と利用時品質に対し現場・ユーザーの立場からレビュー)	2.1.1/2.1.3/2.2.1: 利用時品質の位置づけ・軸・設定方法を確認 (3.3: 機能要件・利用時品質の具体的な例として参照)

1) システムの目的の設定

● 実施事項

事業企画担当

改善したい「現状の課題」を整理、「システムの目的」を設定し、機械学習利用システムが必要か判断

(例えば、配管の点検にかかるコスト削減のために配管の画像診断を行いたい場合は、具体的にどういったコストや課題があるのかなどを検討)

現場担当

課題や目的に対し、現場・ユーザーの立場からレビュー

● ガイドラインの該当箇所

当該項目は、品質保証の入口となる活動だが、具体的な品質保証や信頼性評価の活動の前の段階であるため、ガイドラインの該当箇所はない。

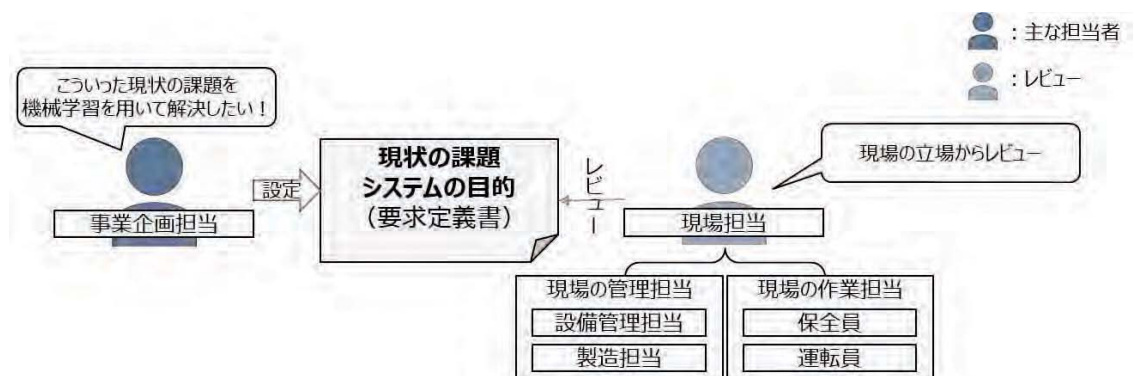


図 4-2 ステップ1「システムの目的の設定」の実施イメージ

2) システムの機能要件・利用時品質の設定

● 実施事項

プラントシステム担当

機械学習利用システムの機能要件と利用時品質を設定

事業企画担当

設定された機械学習利用システムの機能要件と利用時品質に対し企画者の立場からレビュー

現場担当

設定された機械学習利用システムの機能要件と利用時品質に対し現場・ユーザーの立場からレビュー

● ガイドラインの該当箇所

プラントシステム担当は、「2.1.1 利用時品質」「2.1.3 利用時品質・外部品質の軸」で利用時品質の位置づけ、「2.2.1 利用時品質の設定」で、利用時品質の設定の方法を理解する。記載の方法に則って、機械学習利用システムの利用時品質を設定する。事業企画担当・現場担当は、必要に応じて同じ箇所を参照する。

プラントシステム担当は、必要に応じて「3.3 ユースケースに基づく信頼性評価の具体的な適用」のうち、今回構築する機械学習利用システムに類似するユースケースの機能要件・利用時品質の例を参照する。

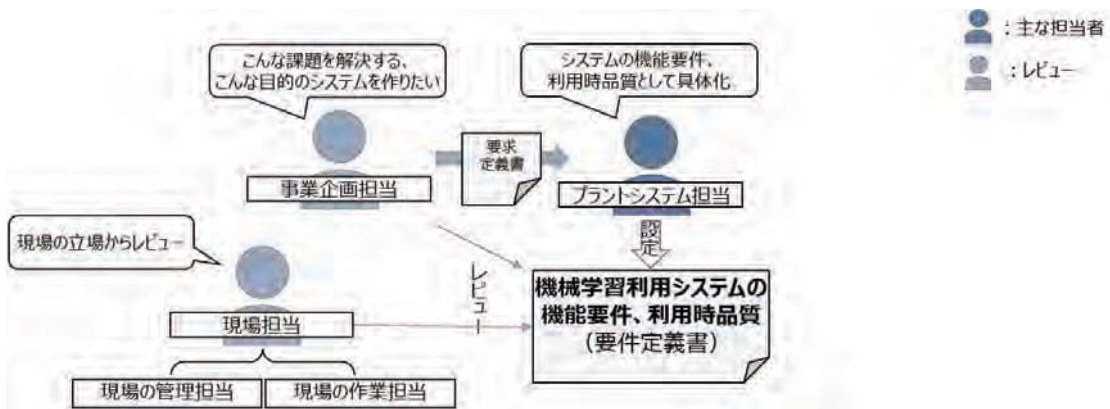


図 4-3 ステップ 2「システムの機能要件・利用時品質の設定」の実施イメージ

(3) 設計

機械学習利用システム全体の設計を行う。すなわち、機械学習利用システム全体に求められる利用時品質に基づき、機械学習要素の外部品質や機械学習要素以外のシステムの役割といった、各構成要素に求められる要件を設定する。

なお、後の「(6)運用」フェーズで実施する各品質の確認のために、本フェーズの時点で、予め「(6)運用」で求められるモニタリングの頻度や実施基準を検討しておくことが重要である。

表 4-9 フェーズ「設計」の実施事項

ステップ	品質保証での実施項目	実施者	ガイドライン該当箇所
3	外部品質の設定	プラントシステム担当: 利用時品質を踏まえ、機械学習要素の外部品質を設定 (現場の管理担当: 設定された外部品質の具体化内容に対し現場の立場からレビュー)	2.1.2/2.1.3/2.2.2: 外部品質の位置づけや設定の方法を確認 (3.3: 外部品質の具体的な例として参照)
4	外部品質のレベルに係る機械学習利用システムと安全関連系の整理	プラントシステム担当: 機械学習利用システムの各構成要素(機械学習要素、機械学習要素以外の構成要素(外部安全機構等))、同システムと独立した安全関連系の有無や機能要件・関係、を整理 (現場の管理担当: 機械学習利用システムの機械学習要素以外の構成要素(外部安全機構等)や同システムと独立した安全関連系について現場の立場からレビュー 環境安全担当: 機械学習利用システムの機械学習要素以外の構成要素(外部安全機構等)や同システムと独立した安全関連系がある場合、その構成要素や安全関連系について安全管理の立場からレビュー)	2.2.3: 機械学習要素以外の構成要素、独立した安全関連系の位置づけに応じた外部品質のレベル設定の考え方を確認
5	外部品質のレベル設定	プラントシステム担当: 機械学習要素の外部品質のレベルの設定 (現場の管理担当: 外部品質のレベルに対し現場の立場からレビュー)	2.2.3: 外部品質のレベル設定の方法を確認

3) 外部品質の設定

- 実施事項

- プラントシステム担当

- 具体化した利用時品質の内容を踏まえ、機械学習要素の外部品質を設定

- 現場の管理担当

- 設定された外部品質に対し現場の立場からレビュー

- ガイドラインの該当箇所

プラントシステム担当は、「2.1.2 外部品質」「2.1.3 利用時品質・外部品質の軸」で、外部品質の位置づけ、「2.2.2 外部品質の設定」で、外部品質の設定の方法を理解する。記載の方法に則り機械学習要素の外部品質の設定をする。現場の管理担当は、必要に応じて同じ箇所を参照する。

プラントシステム担当は、必要に応じて「3.3 ユースケースに基づく信頼性評価の具体的な適用」のうち、今回構築する機械学習利用システムに類似するユースケースの外部品質の具体例を参照する。

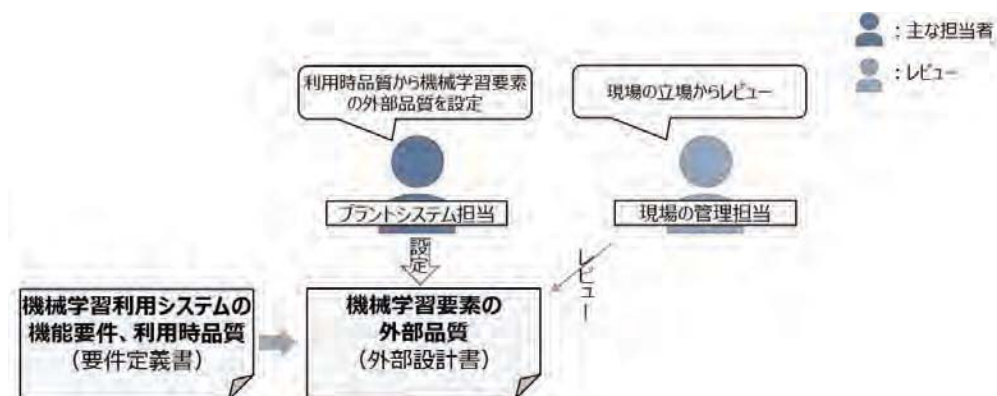


図 4-4 ステップ3「外部品質の設定」の実施イメージ

4) 外部品質のレベルに係る機械学習利用システムと安全関連系の整理

- 実施事項

プラントシステム担当

機械学習利用システムの利用時品質の達成及び機械学習要素の外部品質のレベル設定に必要な同システムの各構成要素（機械学習要素、機械学習要素以外の構成要素（外部安全機構等））、同システムと独立した安全関連系の有無や機能要件を整理

現場の管理担当

機械学習利用システムの機械学習要素以外の構成要素（外部安全機構等）や同システムと独立した安全関連系がある場合、その構成要素や安全関連系について現場の立場からレビュー

環境安全担当

機械学習利用システムの機械学習要素以外の構成要素（外部安全機構等）や同システムと独立した安全関連系がある場合、その構成要素や安全関連系について安全管理の立場からレビュー

- ガイドラインの該当箇所

プラントシステム担当は、「2.2.3 外部品質のレベル設定」で、機械学習利用システムの各構成要素（機械学習要素及び機械学習要素以外の構成要素（外部安全機構等））、また同システムと独立した安全関連系の位置づけに応じた機械学習要素の外部品質のレベル設定の考え方を確認し、同システムの機械学習要素以外の構成要素及び同システムと独立した安全関連系の有無や機能要件を整理する⁶⁶。現場の管理担当・環境安全担当は、必要に応じて同じ箇所を参照する。

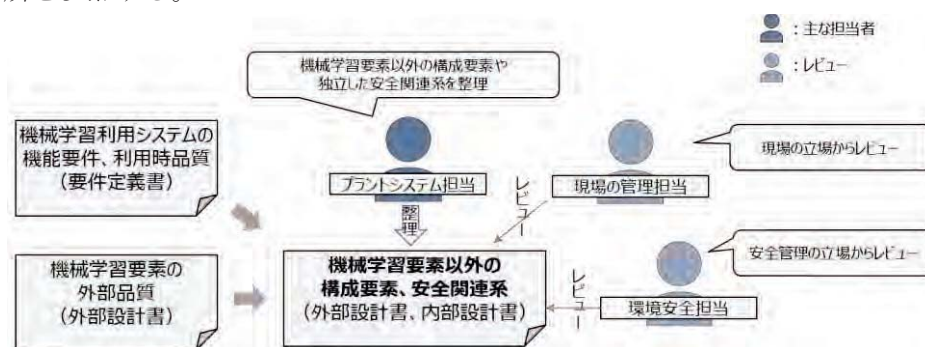


図 4-5 ステップ4「外部品質のレベルに係る機械学習利用システムと安全関連系全体の整理」の実施イメージ

⁶⁶ ここで、機能安全規格 IEC61508(JIS C 0508)、IEC61511(JIS C 0511)等の既存のシステム開発プロセスに基づき、機械学習利用システムと独立した安全関連系により十分に安全が確保されていることが確認できた場合、ステップ 2,3 で設定した利用時品質・外部品質のうち「リスク回避性」軸の項目を機械学習要素に設定する必要はないと判断される。例として「3.3.3 設備劣化診断」のユースケースで「リスク回避性」軸の利用時品質・外部品質を設定しない場合を掲載している。

5) 外部品質のレベル設定

● 実施事項

プラントシステム担当

機械学習要素の外部品質のレベルを設定

現場の管理担当

外部品質のレベルに対し現場の立場からレビュー

● ガイドラインの該当箇所

プラントシステム担当は、「2.2.3 外部品質のレベル設定」で、外部品質のレベルの設定の方法を理解する。機械学習要素の外部品質や、機械学習利用システムの機械学習要素以外の構成要素(外部安全機構等)を踏まえ、記載の方法に則って外部品質のレベル設定を行う。

「リスク回避性」は必要に応じて機械学習利用システム全体の SIL 評価と「AISL 表」の活用を組み合わせ設定する。「パフォーマンス」は基準に基づいて適切なレベルに設定する。

現場の管理担当は、必要に応じて同じ箇所を参照する。

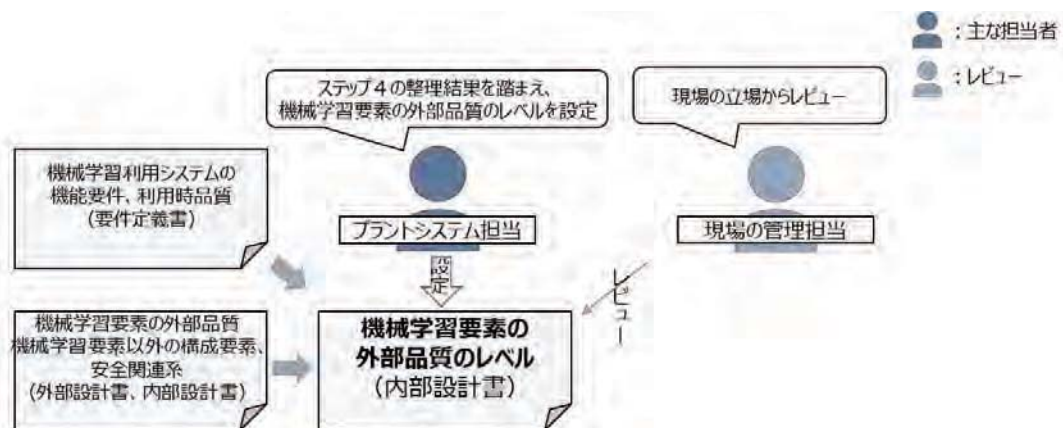


図 4-6 ステップ5「外部品質のレベル設定」の実施イメージ

(4) 実装

機械学習利用システムの実装にあたり、内部品質の実現を通じた機械学習要素の開発、及び機械学習要素以外の構成要素の開発を行う。機械学習要素の信頼性の具体的な確保策を検討し実施する重要な工程である。

表 4-10 フェーズ「実装」の実施事項

ステップ	品質保証での実施項目	実施者	ガイドライン該当箇所
6	内部品質のレベル設定	<u>機械学習設計・開発担当</u> :外部品質のレベルに応じて、機械学習要素の内部品質のレベルを設定	2.1.4:内部品質の8つの軸の概要を理解 2.2.4:内部品質のレベル設定を確認 附録(チェックリスト):内部品質のプラント保安分野での観点を確認 (3.3、附録(チェックリスト):内部品質のユースケース固有の観点を参照)
7	機械学習要素の設計・開発	<u>機械学習設計・開発担当</u> :機械学習要素の設計、内部品質の要求事項と観点に応じた開発 <u>機械学習品質保証担当</u> :機械学習要素の開発において、内部品質の要求事項と観点を満たしているか品質保証の立場から確認 (<u>現場担当・プラントシステム担当</u> :機械学習要素の開発において、データ管理・現場・ユーザーの立場からレビュー、データ提供)	2.2.5、附録(チェックリスト):内部品質の「要求事項」と「観点」の実行についてと具体的に実行する項目を確認
8	機械学習利用システムの機械学習要素以外の構成要素(外部安全機構等)の開発	<u>プラントシステム担当</u> :設定した外部品質・利用時品質に必要な機械学習要素以外の構成要素(外部安全機構等)を開発	—

6) 内部品質のレベル設定

- 実施事項

機械学習設計・開発担当

外部品質のレベルに応じて、機械学習要素の内部品質のレベルを設定し、設定したレベルに対応した要求事項および、プラント保安分野での観点とユースケース固有の観点を確認

- ガイドラインの該当箇所

機械学習設計・開発担当は、「2.1.4 内部品質」で内部品質の8つの軸の内容を、「2.2.4 内部品質のレベルの確認」で、内部品質のレベルの設定方法を理解する。機械学習要素の外部品質及び外部品質のレベルに応じて、機械学習要素の内部品質のレベルを設定する。「附録（チェックリスト）」で、レベル設定に応じた「要求事項」や「プラント保安分野での観点」、また必要に応じて「3.3 ユースケースに基づく信頼性評価の具体的適用」のうち、今回構築する機械学習利用システムに類似するユースケースの「ユースケース固有の観点」を確認する。プラントシステム担当は、必要に応じて同じ箇所を参照する。

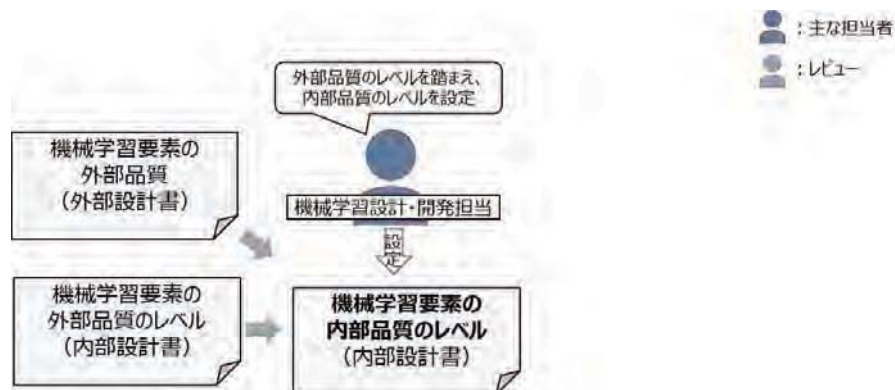


図 4-7 ステップ6「内部品質のレベル設定」の実施イメージ

7) 機械学習要素の設計・開発

● 実施事項

機械学習設計・開発担当

機械学習要素の設計（機械学習要素の具体的な仕様やモデルの特定）、内部品質の要求事項と観点に従った開発

機械学習品質保証担当

機械学習要素の開発において、内部品質の要求事項と観点を満たしているか品質保証の立場から確認

現場担当・プラントシステム担当

機械学習要素の開発において、データ管理・現場・ユーザーの立場からレビューやデータ提供、ラベル付けに協力)

● ガイドラインの該当箇所

機械学習設計・開発担当は、「2.2.5 内部品質の要求事項の確認および実行」で「要求事項」や「観点」の位置づけや実行方法について確認し、ステップ6で確認した「附録（チェックリスト）」における「要求事項」と「観点」に応じて、機械学習要素の設計・開発を行う。またそのプロセスにおいて、内部品質の「要求事項」や「観点」が満たされているか、機械学習品質保証担当がガイドライン及び「附録（チェックリスト）」の同様の箇所を踏まえて確認し、対応の記録を保存する。現場担当・プラントシステム担当は、必要に応じて同じ箇所を参照する。

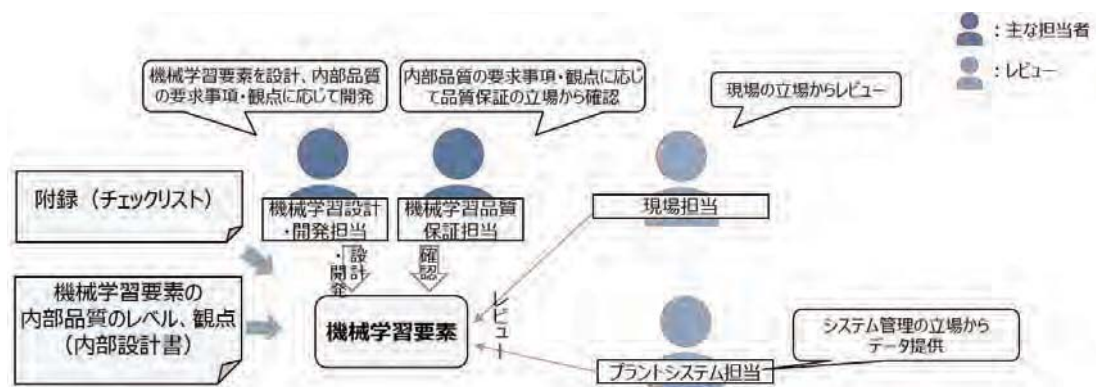


図 4-8 ステップ7「機械学習要素の設計・開発」の実施イメージ

8) 機械学習利用システムの機械学習要素以外の構成要素（外部安全機構等）の開発

- 実施事項

プラントシステム担当

設定した利用時品質および外部品質を満たすために必要な機械学習利用システムの機械学習要素以外の構成要素（外部安全機構等）を開発

- ガイドラインの該当箇所

当該項目は、機械学習要素の外部品質のレベルの前提として必要な要件を満たし、利用時品質を達成するために必要である。しかし、機械学習要素には直接関わらない項目であるため、ガイドラインの該当箇所はない⁶⁷。

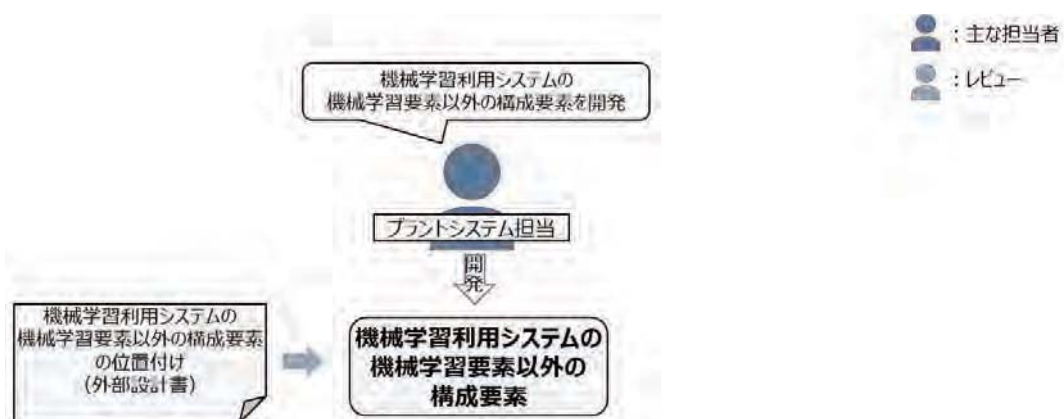


図 4-9 ステップ 8「機械学習要素以外の構成要素（外部安全機構等）の開発」の実施イメージ

⁶⁷ 機械学習要素以外の構成要素（外部安全機構等）は、機能安全規格 IEC61508(JIS C 0508)、IEC61511(JIS C 0511)等の既存のシステム開発プロセスに従って開発する。

(5) テスト・検収

機械学習利用システムをテストし、その結果を評価する。また、テストの結果も踏まえ検収を行う。機械学習要素の内部品質は、ステップ9「機械学習要素の設計・開発」の中で確認されるので、本フェーズにおいては、機械学習利用システムの中における機械学習要素の外部品質の評価を中心に行う。検収では、利用時品質を含む要求事項が確認される。

表 4-11 フェーズ「テスト・検収」の実施事項

ステップ	品質保証での実施項目	実施者	ガイドライン該当箇所
9	テスト(外部品質の評価)	<u>システム品質保証担当</u> :機械学習利用システムのテストを行い、その結果を評価。機械学習要素の外部品質が要求水準に達しているか評価 (<u>プラントシステム担当</u> :テスト結果の評価を踏まえて、外部安全機構・安全関連系の整理や、外部品質の設定及びレベル設定をした立場からレビュー <u>環境安全担当</u> :テスト結果の評価に対し、安全管理の立場からレビュー)	2.2.3:外部品質の評価基準を確認
10	検収	<u>システム品質保証担当</u> :テストでの外部品質の評価結果から機械学習利用システム全体の評価を行い、それらの評価結果等を踏まえ基準を満たす機械学習利用システムであれば検収 (<u>事業企画担当</u>:機械学習利用システム全体の評価をプロジェクト発案者の立場からレビュー <u>現場担当</u>:利用時品質の評価結果に対し現場・ユーザーの立場からレビュー)	—

9) テスト（外部品質の評価）

● 実施事項

システム品質保証担当

機械学習利用システムのテストを行い、その結果を評価。機械学習要素の外部品質が要求水準に達しているか評価

プラントシステム担当

テスト結果の評価に対し、外部安全機構・独立した安全関連系の整理や、外部品質及びそのレベル設定をした立場からレビュー

環境安全担当

テスト結果の評価に対し、安全管理の立場からレビュー

● ガイドラインの該当箇所

システム品質保証担当は、「2.2.3 外部品質のレベル設定」で、外部品質の評価基準と方法を理解する。機械学習要素の外部品質の設定内容やレベルを踏まえ、テスト結果のうち機械学習要素の外部品質が要求水準に達しているか評価する。プラントシステム担当・環境安全担当は、必要に応じて同じ箇所を参照する。

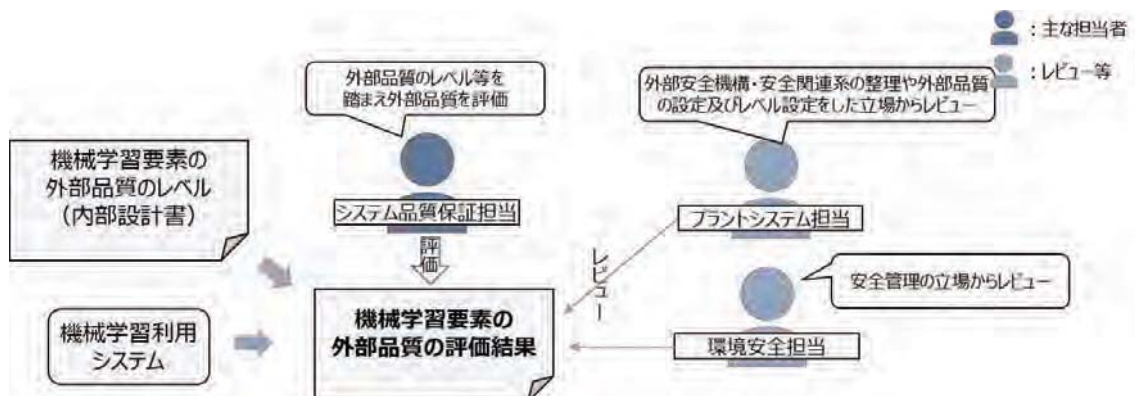


図 4-10 ステップ9「テスト（外部品質の評価）」の実施イメージ

10) 検収

● 実施事項

システム品質保証担当

テストでの外部品質の評価結果から機械学習利用システム全体の評価を行い、それらの評価結果等を踏まえ基準を満たす機械学習利用システムであれば検収

事業企画担当

機械学習利用システム全体の評価をプロジェクト発案者の立場からレビュー

現場担当

利用時品質の評価結果に対し現場・ユーザーの立場からレビュー

● ガイドラインの該当箇所

検収は各社の基準に基づいて行われるため、ガイドラインで直接該当する箇所はないが、外部品質のレベルや、それに応じて設定した内部品質のレベル、内部品質の要求事項および対応の記録、またテストでの外部品質の評価結果を適切に踏まえる必要がある。

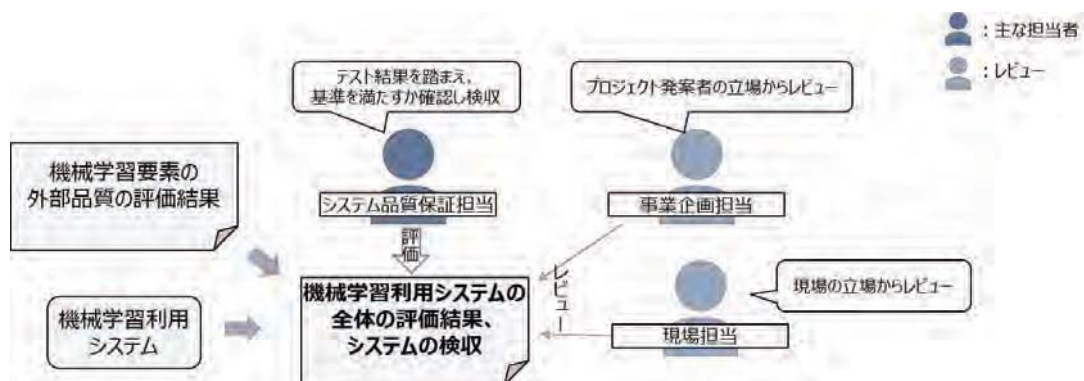


図 4-11 ステップ 10「検収」の実施イメージ

(6) 運用

機械学習利用システムの運用開始後は、必要に応じてシステムの更新が行われる。品質検査を都度経て更新する場合、リアルタイムで更新する場合のいずれにおいても、機械学習要素及び機械学習利用システム全体のアウトプット（利用時品質・外部品質）のモニタリングを行う必要がある。また、その結果を踏まえシステムの更新をする必要があるが、その判断の基準などの仕組みは、運用開始前（「(3)設計」フェーズの時点）に用意することが必要な場合が多い⁶⁸。

上記を踏まえ、運用フェーズの実施項目は、予め設定した頻度や実施基準に基づいて実施する品質保証の活動について記載する。

表 4-12 フェーズ「運用」の実施事項

ステップ	品質保証での実施項目	実施者	ガイドライン該当箇所
11	利用時品質の確認	システム品質保証担当 :運用中の機械学習利用システムの利用時品質を確認 (現場担当:利用時品質の確認結果に対し現場・ユーザーの立場からレビュー)	2.1.1./2.1.3/2.2.1:利用時品質の概要を確認
12	外部品質の確認	システム品質保証担当 :運用中の機械学習利用システムの利用時品質の確認結果を踏まえ、外部品質を確認 (プラントシステム担当 :外部品質の確認結果を、外部安全機構・安全関連系の整理や、外部品質の設定及びレベル設定をした立場からレビュー)	2.2.3:外部品質の評価基準を確認
13	内部品質(運用時品質の維持性)の確認	機械学習品質保証担当 :運用中の機械学習利用システムの利用時品質・外部品質の確認結果も踏まえ、内部品質の要求事項の対応状況を確認	2.2.4:内部品質の評価基準を確認 2.2.5、附録(チェックリスト):内部品質の要求事項と観点を確認
14	システムの更新	機械学習設計・開発担当 :外部品質・内部品質の確認結果に応じて機械学習要素を更新 プラントシステム担当 :機械学習要素以外の構成要素(外部安全機構等)を更新、システム管理の立場から機械学習要素の更新に対しデータ提供 (現場担当:機械学習要素の更新において、現場・ユーザーの立場からレビュー)	2.2.4:内部品質の評価基準を確認 2.2.5、附録(チェックリスト):内部品質の要求事項と観点を参照

⁶⁸ 産業技術総合研究所（2020）「機械学習品質マネジメントガイドライン 第1版」

11) 利用時品質の確認

- 実施事項

システム品質保証担当

運用中の機械学習利用システムの利用時品質の達成状況を、設定したタイミングで確認

現場担当

利用時品質の確認結果に対し現場・ユーザーの立場からレビュー

- ガイドラインの該当箇所

システム品質保証担当は、「2.1.1 利用時品質」「2.1.3 利用時品質・外部品質の軸」で利用時品質の位置づけ、「2.2.1 利用時品質の設定」で、利用時品質の設定の方法を理解する。構築時のシステムの機能要件・利用時品質等を踏まえ、運用中のシステムの利用時品質が当初の目的を達成できているか確認する。現場担当は、必要に応じて同じ箇所を参照する。

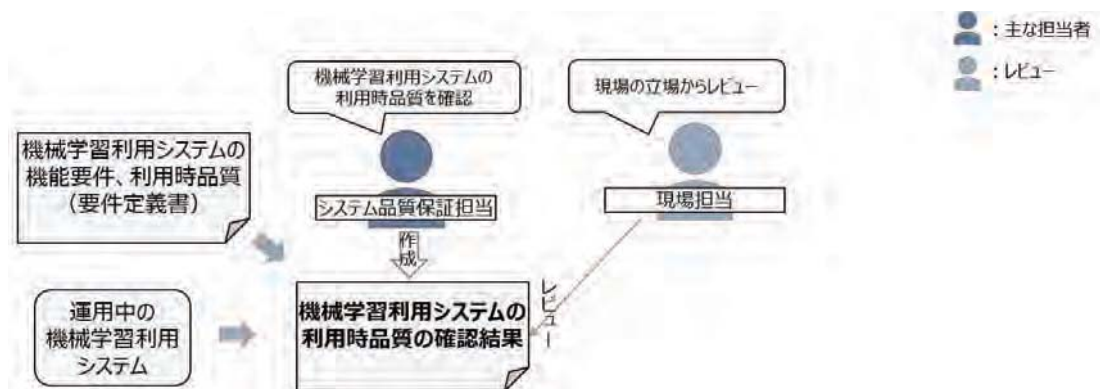


図 4-12 ステップ 11「利用時品質の確認」の実施イメージ

12) 外部品質の確認

● 実施事項

システム品質保証担当

運用中の機械学習利用システムの利用時品質の確認結果も踏まえ、機械学習要素の外部品質の達成状況を確認

プラントシステム担当

外部品質の確認結果を、外部安全機構・独立した安全関連系の整理や、外部品質及びそのレベル設定をした立場からレビュー

● ガイドラインの該当箇所

システム品質保証担当は、「2.2.3 外部品質のレベル設定」で、外部品質の評価基準と方法を理解する。機械学習利用システムの機能要件や利用時品質、機械学習要素の外部品質の設定・レベル設定を踏まえ、運用中の機械学習要素の外部品質が当初設定した要求事項を達成できているか確認する。プラントシステム担当は、必要に応じて同じ箇所を参照する。

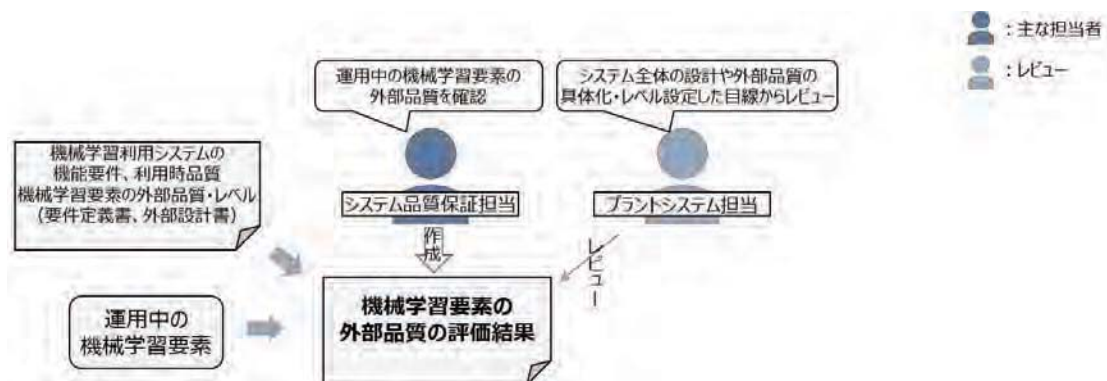


図 4-13 ステップ 12「外部品質の確認」の実施イメージ

13) 内部品質（運用時品質の維持性）の確認

- 実施事項

機械学習品質保証担当

運用中のシステムの利用時品質・外部品質の確認結果も踏まえ、「運用時品質の維持性」⁶⁹の内部品質の要求事項の対応状況を確認

- ガイドラインの該当箇所

機械学習品質保証担当は、「2.2.4 内部品質のレベルの確認」及び「2.2.5 内部品質の要求事項の確認および実行」「附録（チェックリスト）」を踏まえ、内部品質の「要求事項」と「観点」に対応できているか、特に「運用時品質の維持性」について確認する。



図 4-14 ステップ 13「内部品質の確認」の実施イメージ

⁶⁹ 「運用時品質の維持性」の内部品質を満たすことを通して、他の内部品質の達成状況が低下していないかを確認する。

14) システムの更新⁷⁰

● 実施事項

機械学習設計・開発担当

外部品質・内部品質の確認結果に応じて機械学習要素を更新

プラントシステム担当

利用時品質・外部品質の確認結果に応じて機械学習要素以外の構成要素(外部安全機構等)を更新、システム管理の立場から機械学習要素の設計・開発に対しデータ提供

現場担当

機械学習要素の更新において、現場・ユーザーの立場から必要に応じて意見

● ガイドラインの該当箇所

機械学習設計・開発担当は、「2.2.4 内部品質のレベルの確認」及び「2.2.5 内部品質の要求事項の確認および実行」「附録(チェックリスト)」を踏まえた確認結果に応じて内部品質に修正を加えて、機械学習要素を更新する。現場担当・プラントシステム担当は、必要に応じて同じ箇所を参照する。機械学習要素以外の更新についてはガイドラインでの該当箇所はない。

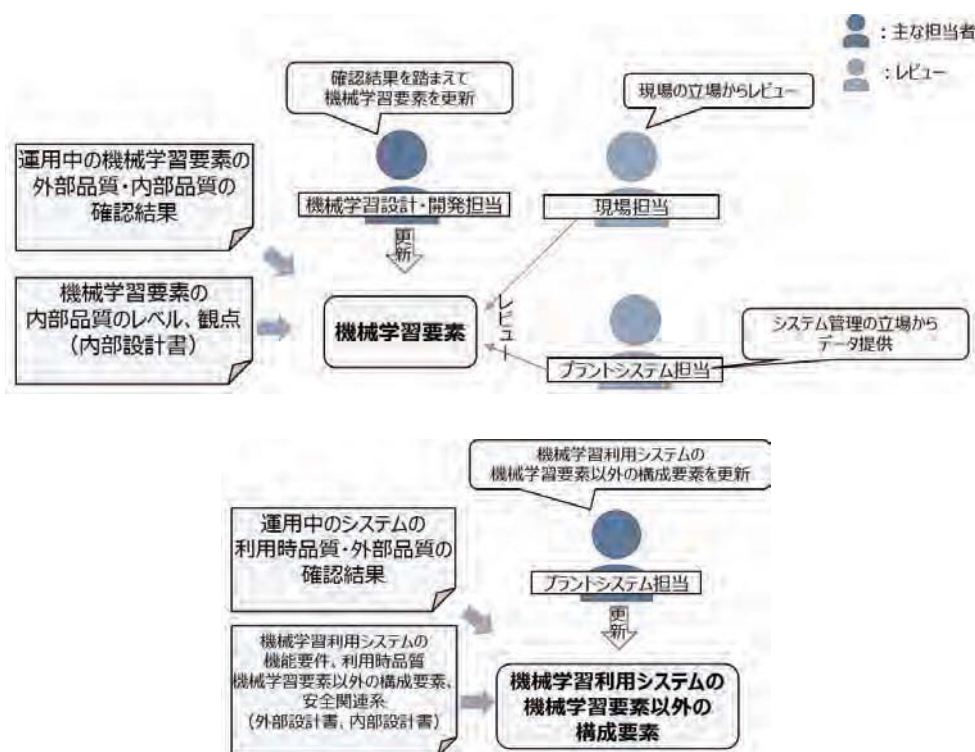


図 4-15 ステップ 14「システムの更新」の実施イメージ

⁷⁰ 更新したシステムの外部品質は、テスト・検収フェーズに戻り、ステップ 9 (テスト) で確認する。

附録

内部品質確保のための「プラント保安分野での観点」チェックリスト

- 0 (チェックリストを見る前に) 機械学習要素の開発にあたり、要求されている内部品質の要求レベル (レベル1～3) を確認する。(ガイドライン2.2.4項)
- 1 下記「本チェックリストの使用場面」に応じて検討する内部品質軸について、該当する要求レベルの要求事項を確認する。
- 2 該当する要求レベルの要求事項と直接関連する「プラント保安分野での観点」を確認する。「内部品質の要求レベル」欄で該当する要求レベル (Lv1～Lv3) 開発する機械学習要素がユースケースと類似している場合は、同様に「ユースケース固有の観点」も確認する。
- 3 要求事項を問わない共通的な「プラント保安分野での観点」を確認する。「内部品質の要求レベル」欄には、「共通」と記載している。
開発する機械学習要素がユースケースと類似している場合は、同様に「ユースケース固有の観点」も確認する。

【使用場面のリスト】

- # 1「データの設計」の検討(要求分析の十分性、データ設計の十分性)

検討の例：

- ・データ収集でカバースべき属性、環境要因等の検討
- ・要求分析や収集データについての現場のエンジニアへの確認等

- ## 2 「データの品質」の検討（データセットの被覆性、データセットの均一性）

検討の例：

- ・カバすべき属性・環境要因等が収集できていることの確認
- ・データ不足時のシミュレーション活用等の検討等

- ### 3 「学習済みモデルの品質」の検討（機械学習モデルの正確性、機械学習モデルの安定性）

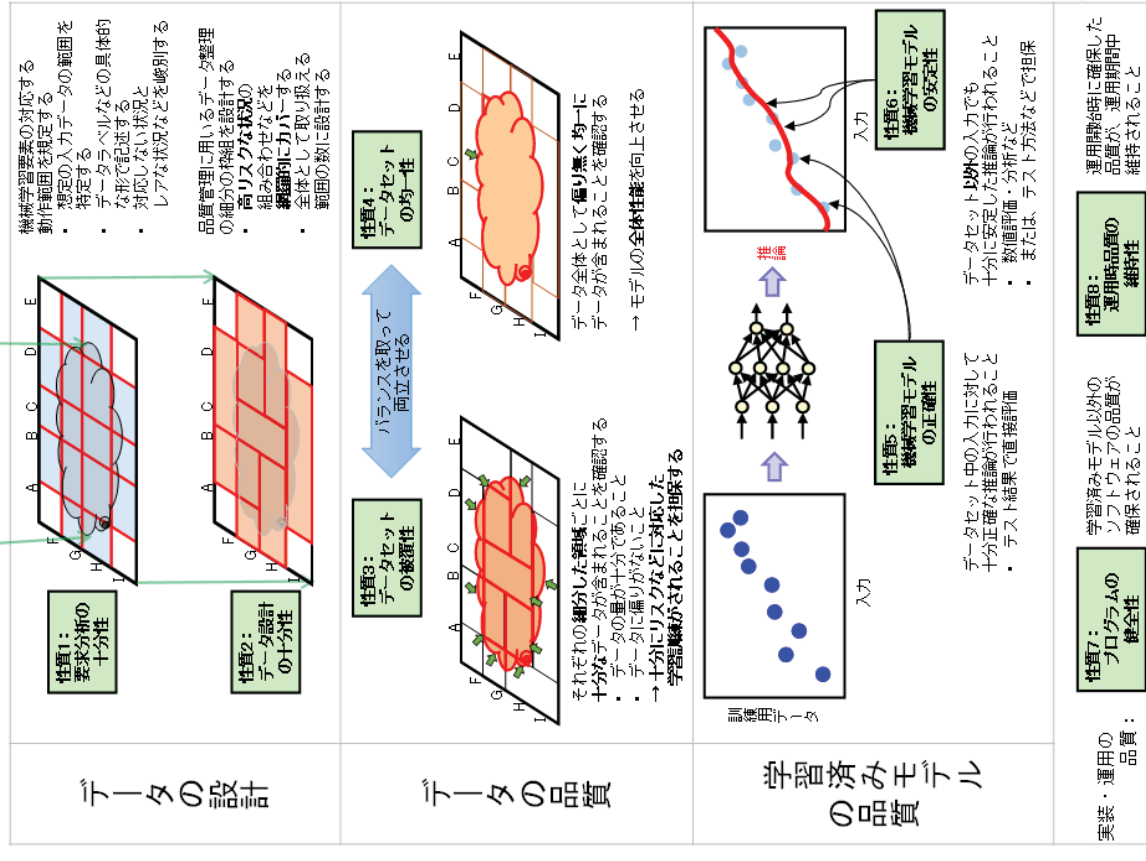
検討の例：

- ・誤判断の許容水準の検討
- ・テスト手法の検討等

- #### 4「実装・運用の品質」の検討（プログラムの健全性、運用時品質の維持性）

検討の例：

- ・実装に利用するソフトウェアの検討
- ・精度検証・再学習の頻度設計 等



実装	運用の品質：	性質7： 学習済みモデル以外のソフトウェアの品質が確保されること 健全性	性質8： 運用時品質の維持されること 維持性	運用開始時に確保した品質が、運用期間中維持されること 維持性
----	--------	--	------------------------------	-----------------------------------

[illegible]

[illegible]

1		Lv1	利用するソフトウェアについては、信頼できる実績を持つソフトウェア等を選定し、その選定経緯を記録すること。	・シミュレータを活用する場合、実績を確認し、その選定経緯を記録しましたか？	－	－	－	－	－	－
1		Lv1	選定したソフトウェアについて、その欠陥の発見等を運用期間中モニタリングし、必要に応じて修正等の措置をとること。	・シミュレータを活用する場合、欠陥のモニタリングや修正は可能ですか？	－	－	－	－	－	－
1		Lv1	学習からテストフェーズに至るまでの環境と、実用段階で用いる環境の相違について、その影響などをあらかじめ検討しておくこと。	－	－	－	－	－	－	－
2		Lv2	利用するソフトウェアについて、検査・実験等によりその信頼性を自己評価すること。	・シミュレータを活用する場合、検査・実験等で信頼性を自己評価しましたか？	－	－	－	－	－	－
2		Lv2	可能な場合には、SIL1相当のソフトウェア信頼性を得られたソフトウェアを用いること。	－	－	－	－	－	－	－
2		Lv2	システムの運用期間中のソフトウェアの健全性の維持に關する保守体制を必ず構築すること。	・シミュレータを活用する場合、保守体制が整っていますか？	－	－	－	－	－	－
2		Lv2	バリデーションおよびテストフェーズにおいて、原則として実用段階で用いられる計算環境（浮動小数点精度・モデル規模等）を模倣した環境でバリデーション・テストを行うこと。または、テスト済み学習モデルと実用環境での学習モデルの動作の一致性について、何らかの検証を行う事。	－	－	－	－	－	－	－
		Lv3	SIL 1（またはシステムの要求するSILレベル）のソフトウェア品質の確認を必ず行うこと。	－	－	－	－	－	－	－
		Lv3	実用環境の計算環境での学習モデルの振る舞いに基づくテスト（または形式検証など）を必ず行うこと。	－	－	－	－	－	－	－
		Lv3	また、そのモデルと実用環境での動作の一致の確認を、結合テスト以降の段階で必ず行うこと。	－	－	－	－	－	－	－
		共通	・以下のような場合に、機械学習要素の計算量を適切に抑えられるよう、考慮しましたか？ - 特殊な環境・デバイスで、計算リソースが限られる場合（エッジデバイス、特殊なPC等で計算する場合） - 同じデバイス上で他の処理も動いており、それらの処理への影響が許されない場合 等 ・システムを評価するときに、外部ライブラリに対する単体テストやシステムテストを実施していますか？	－	－	－	－	－	－	－
		共通	・ライブラリのサブライヤとの間で、不具合に対する責任範囲は明確になっていますか？	－	－	－	－	－	－	－
		共通	・OSやOSS等の各種ソフトウェアの更新頻度やサポート期間を考慮してソフトウェアを利用していますか？ソフトウェアのアップデートに対する対応や、サポートが終わった際の対応について決めていますか？	－	－	－	－	－	－	－
		共通	・特にセキュリティアップデートがある場合など、OSやOSS等のソフトウェアのアップデートに対してシステムの更新を行っているか	－	－	－	－	－	－	－
1		Lv1	外部環境変化によりシステムの品質が著しく失われたときに対応について、あらかじめ検討しておくこと。	・プラント分野における「外部環境変化」を抽出しましたか？ - プラントの修繕や経年劣化、運転条件の変化など	－	－	－	－	－	－
1		Lv1	オンラインで追加学習を行う場合には、予想外の品質の低下がもたらす影響についてあらかじめ検討しておくき、必要な場合には動作範囲の限定などのシステムの的な対応を取ること。	－	－	－	－	－	－	－
1		Lv1	オフラインで追加学習を行う場合には、前7項に準じた品質管理を行う事。	－	－	－	－	－	－	－
2		Lv2	製品の利用状況が許す範囲において、システムの品質について、動作結果との対象などから品質劣化・誤判断のモニタリングを行うこと。モニタリングにおいては、ブライバシーなど製品品質以外の要因を十分に検討すること。	・プラント分野における「製品品質以外の要因」を抽出しましたか？ - プラント事業者からAIベンダーへのデータ開示の度合いなど	－	－	－	－	－	－
			オンラインでの追加学習を行う場合には、追加学習結果を何らかの方法で定常的にモニタリングす	－	－	－	－	－	－	－

[illegible]

プラントにおける AI の信頼性評価に関する検討会 委員等名簿

座長(敬称略)

山下善之 国立大学法人 東京農工大学大学院工学研究院応用化学部門 教授

委員(五十音順、敬称略)

青山敦 立命館大学 テクノロジー・マネジメント研究科 教授
 石川冬樹 国立情報学研究所 アーキテクチャ科学研究系 准教授
 江崎宣雄 三井化学株式会社 生産・技術本部 生産技術高度化推進室長
 大岩寛 国立研究開発法人産業技術総合研究所 サイバーフィジカルセキュリティ研究センター ソフトウェア品質保証研究チーム長
 大野拓也 日揮グローバル株式会社 技術イノベーションセンター 特別理事
 岡田義昭 日本電気株式会社 コーポレート事業開発本部 上席プロフェッショナル
 小山田賢治 高圧ガス保安協会 高圧ガス部長代理
 国頭庸一 三菱ケミカル株式会社 生産技術部 安全工学センター長
 小淵恵一郎 横河電機株式会社 IA プロダクト&サービス事業本部 インフォメーションテクノロジーセンター AI ビジネス開拓部 部長
 近藤晃弘 株式会社日立製作所 産業・流通ビジネスユニット ソリューション&サービス事業部 産業 PA ソリューション部
 齋藤剛 独立行政法人労働者健康安全機構 労働安全衛生総合研究所 機械システム安全研究グループ 上席研究員
 喰田秀樹 出光興産株式会社 生産技術センター システム高度化技術室長
 西 晴樹 消防研究センター 火災災害調査部長
 濱田聡之 昭和電工株式会社 レスポンスブルケア部 環境安全室
 原田典明 旭化成株式会社 理事 生産技術本部 デジタルイノベーションセンター長
 藤井達也 ENEOS 株式会社 工務部設備管理グループマネージャー
 榎谷昌隆 JSR 株式会社 生産技術部長
 丸山宏 株式会社 Preferred Networks PFN フェロー
 安井威公 千代田化工建設株式会社 デジタルトランスフォーメーション本部 デジタル企画マーケティング部 専門長

オブザーバー (敬称略)

石油連盟
 石油化学工業協会
 一般社団法人 日本化学工業協会
 一般財団法人 エンジニアリング協会
 一般社団法人 日本電気計測器工業会
 一般社団法人 日本ディープラーニング協会
 日本メンテナンス工業会
 独立行政法人 経済産業研究所
 神奈川県 くらし安全防災局防災部 消防保安課
 三重県 雇用経済部 ものづくり産業振興課
 四日市市 商工農水部 商工課

総務省消防庁 特殊災害室
 総務省消防庁 危険物保安室
 厚生労働省 労働基準局 安全課
 厚生労働省 労働基準局 化学物質対策課
 経済産業省 製造産業局 素材産業課
 経済産業省 商務情報政策局 情報経済課
 経済産業省 産業保安グループ 高圧ガス保安室
 経済産業省 資源エネルギー庁 資源・燃料部 石油精製備蓄課

事務局

株式会社三菱総合研究所

プラント保安分野 AI 信頼性評価ガイドライン

2020 年 11 月

石油コンビナート等災害防止 3 省連絡会議
(経済産業省、総務省消防庁、厚生労働省)

「プラントにおける AI の信頼性評価に関する検討会」
事務局 三菱総合研究所