

## 故障予測及び作業内容予測機能を持つ 保守作業支援システムの実現可能性の検証

### Verification of the Feasibility of a Maintenance Support System with Failure Prediction and Task Forecasting Functions

中村 伊吹/Ibuki NAKAMURA<sup>1</sup>・左座 祐之助/Yunosuke ZOUZA<sup>1</sup>・広瀬 啓雄/Hiroo HIROSE<sup>1</sup>

・橋本 幸二郎/Kohjiro HASHIMOTO<sup>1</sup>・三代沢 正/Tadashi MIYOSAWA<sup>1</sup>

<sup>1</sup>公立諏訪東京理科大学大学院

#### [Abstract]

To verify the feasibility of a maintenance work support system that reduces long-term downtime and technology transfer costs in maintenance work, we will construct a parts replacement prediction model and a work content prediction model that will serve as the core of the system. The parts replacement prediction model learned 11 ink head replacement records from the equipment usage history, and detected 3 out of 6 replacements within 30 days of the replacement date in the evaluation data. In addition, the work content prediction model predicted work categories based on the records of phenomena when problems occurred, with a correct answer rate of 82.3%. One issue with the prediction is that the parts replacement prediction model does not have a constant detection period for head replacement, which may lead to unnecessary parts replacement. The work content prediction model is not able to identify the causes of failures and the areas that need to be treated. Based on the above issues, we will improve the detection accuracy by improving the data acquisition interval and identify the causes of failures by making the records of problem occurrences into category variables, and work toward the construction and evaluation of the system.

#### [キーワード]

保守作業支援、機械学習、大判インクジェットプリンタ、予知保全、作業内容推定

## 1. はじめに

A 社では業務用大判インクジェットプリンタの製造販売と保守サービスを展開している。大判インクジェットプリンタは、顧客が販売する製品のシールやラベル作成等の業務に利用されている。大判インクジェットプリンタに問題が発生した場合、顧客の業務に影響を及ぼすため、A 社での保守業務は顧客の生産性において重要な役割を担っている。A 社の保守業務では、顧客から問題発生の通報を受け、修理等の必要があった場合、顧客のもとに保守作業を送り、復旧作業を行っている(図 1)。現状の保守業務では、長時間のダウンタイム発生と技術伝承コストの増加が問題視されている。

### (1) 長時間のダウンタイム発生について

機器のダウンタイムは顧客業務の生産性を低下させる可能性があるため、最小限に収める必要がある。しかし、A 社の保守業務では、顧客から問題発生の通報を受けて、保守作業に取り掛かるまでに通常 1~3 日経過してしまう。加えて保守作業が他の作業に当たっていたり、必要な交換部品が不足していたりした場合、保守業務に取り掛かるまでに 7~10 日程経過してしまうことがある。問題発生から保守業務が行われるまでの期間、顧客は機器を利用した業務が出来なくなり、業務の生産性が低下する恐れがある。以上のことから、長時間のダウンタイムが発生している要因として、保守作業に必要な人員や交換部品の手配に多くの時間がかかることが挙げられる。

一案として、問題発生にかかわらず定期的な点検やメンテナンスを行う方法が考えられるが、これでは現状において人員が不足しているにも関わらず、より多くの人員を確保する必要があるため、本研究では対象としない。一方で、問題発生の予兆をセンサー等の情報から捉えることが出来れば、問題発生を A 社で把握できるようになり、計画的に人員や部品の手配を行うことが期待される。よって、本研究で構築を目指すシステムには、問題発生の予兆を捉えることで、計画的な保守業務を遂行するための機能が求められている。

### (2) 技術伝承コスト増加について

技術伝承にかかるコストは日々増加しており、保守作業者の減少がその要因として挙げられる。保守に必要な作業の判断には、機器の現象から故障の原因を複合的に考慮する必要があり、作業者の経験や知識等のノウハウが重要になる。そのため、保守作業技術の習得にはマニュアル化された知識を学習するだけでなく、熟練作業者のノウハウを学ぶことで、保守作業に必要な技術を短期間で習得することが可能になる。

しかし、作業者のノウハウを学習するためにはいくつか問題がある。作業者のノウハウには、言語では表現しきれない非言語的な要素が多く、マニュアル化が困難である。仮にノウハウを文書や動画で蓄えることが出来たとしても、大量のデータからマニュアル作成に有効なデータだけを抽出し、活用することは困難である。現在、製造業の現場では人材不足が深刻化しており、特に保守作業者のような技能人材の確保に課題を抱えている企業が増加している[1]。その中で、熟練作業者がセミナー形式でノウハウを伝える方法では、多くの時間がかかってしまい、熟練作業者が現場で働く時間を減少させてしまう問題がある。このように作業者のノウハウを共有するには多くの課題が挙げられる。

A社では保守作業記録が蓄積されている。保守作業記録には問題発生時の現象、原因、対策が記載されており、これらは保守作業の際に得られたノウハウである。前述の通り、保守作業記録から必要なデータのみを抽出してマニュアル化することは困難である。とはいえ、問題発生時の現象から保守に必要な故障原因や作業内容を予測することが出来れば、熟練作業者のノウハウを別の作業者が利用できるようになり、技術伝承にかかるコストの削減が期待される。したがって、本研究で構築を目指すシステムには、効率的にノウハウ共有を行うための作業内容を予測する機能が求められる。

以上より本研究の目的は、長時間のダウンタイムと技術伝承コストを削減する保守作業支援システムの実現可能性の検証として、システムの中核となる2つの予測モデルの構築を行う。

保守作業支援システムには、主に2つの機能がある。1つ目は計画的なリソース確保を行うための故障予測機能、2つ目は効率的にノウハウ共有を行うための作業内容予測機能である。本研究ではこの2つの機能の中心である故障予測モデルの一部と作業内容予測モデルの構築を行う。

以下、2章で先行研究の紹介と本研究との違い、3章で構築を目指す保守作業支援システムの構成について述べ、4章で部品交換予測モデルの構築、5章で作業内容予測モデルの構築、6章で本研究のまとめについて述べる。

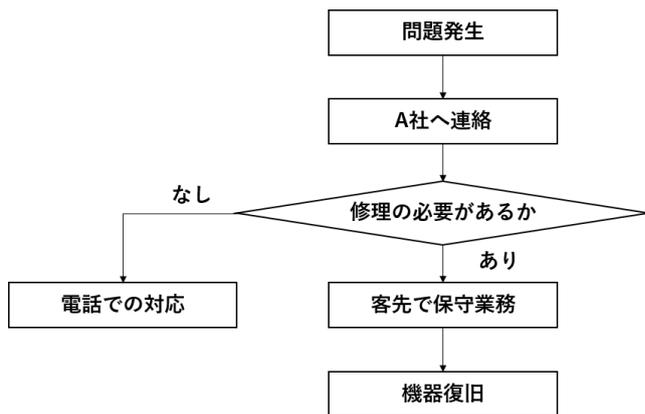


図- 1 A社における保守対応

## 2. 先行研究

近年の製造業における予知保全ではIoT技術の発展に伴い、工場設備や機器に設置されたセンサーデータに対して機械学習を用いて分析する方法が行われている[2]。データの分析に機械学習を用いる利点として、データの解析や予測において、専門的な知識を必要としないことが挙げられる[3][4]。最近では Automated Machine Learning と呼ばれる、機械学習のモデル構築を自動化するサービスも提供されており、機械学習の活用が容易になってきている[5]。これは、多くの企業にとって、データ分析への敷居を下げ、新たなシステムやサービスの提供に繋がる。

機械学習を用いた故障予測を予知保全へ活用しようとする研究も報告されている。宇宙機の故障予兆検知では、教師なし学習の異常検知を用いて X 線天文衛星の電源系機器の故障予兆検知を試み、一定条件下で故障の予兆を検出することが出来ている[6]。教師あり学習のクラス分類を用いた工場機器の故障予測では、化学プラントに取り付けられたセンサーデータからロジスティック回帰、SVM、ニューラルネットを用いてモデル作成、評価実験を

行い、発生頻度の高い故障に対して高い予測精度を得ている[7]。これらの先行研究では発生頻度が高い故障への予測や特定条件下での予測には成功しているものの、発生頻度が低い故障の予測を行えていない問題を抱えている。

他にも、本研究で扱うプリンター機器の保守作業を支援するシステムへの取り組みも行われている。[8]のリコーでは全世界の200の国と地域にMFP (Multifunction Printer) の販売・サービスを展開しており、日本国内だけでも400か所のサービス拠点を有し、数千名の保守担当者がMFPの保守サポートを行っている。リコーでは蓄積されたビッグデータを解析することで、故障の予兆検知と検知後のアクションを定めたシステム開発を行い、保守作業の現場での活用を実現させている。

教師あり学習を用いた故障予測に関する先行研究では、発生頻度が低い故障の予測を行えない問題点がある。そこで本研究では、発生頻度の高い故障を教師あり学習のクラス分類を用いて予測し、発生頻度の低い故障については教師なし学習の異常検知を用いる。私の知る限り、先行研究ではこの2種類の予測モデルを利用した保守作業支援システムに関する研究はない。

次に[8]の先行研究では故障の検出から、現場に必要な作業の確定までを行っている。この先行研究と比較して本研究では故障に関するビッグデータを蓄積できず、同様の予測方法では困難である。そこで、本研究では保守作業記録から保守に必要な作業内容を予測することで、故障検出後の保守業務支援を行う。

### 3. システムの構成

1章で述べたようにA社では、保守業務において計画的に人員や交換部品の調達を行うことと、作業者のノウハウを効率的に活用することが出来ていない。そこで、本研究で構築する保守作業支援システムでは故障予測機能と作業内容予測機能を作成することで、計画的な保守業務の遂行と技術伝承コストの削減に取り組む。それぞれの機能の目的と目標について表1に示す。

故障予測の目標として、故障が発生する30日以内に検出することを目標に設定した。これは、故障の検出が早すぎる場合には、本来稼働できたはずの時間を無駄にしまう恐れがあるためである。作業内容予測での目標は、作業内容予測を活用した保守作業を行うことで、保守業務にかかる時間を50%削減することが出来れば、熟練作業者のノウハウを効率的に活用できているとして設定した。構築する保守作業支援システムの手順は次の通りである。また、システムフローを図2に示す。

- (1) 通常運転している機器から予測に必要なデータを取得
- (2) 取得したデータから部品交換の予測と機器の異常度を計算
- (3) 各予測モデルから得られた値が一定の閾値を超えた場合、保守に必要な人員と交換部品の確認と手配
- (4) 顧客に問題発生の有無を確認
- (5) 問題が発生していた場合、顧客のもとへ出向く
- (6) 保守作業予測を参考に保守作業を行う
- (7) 機器復旧

表-1 保守作業支援システムの機能の目的と目標

機能	目的	目標	背景
故障予測	計画的な人員や交換部品の手配	故障が発生する30日以内の予測	長時間のダウンタイム発生による顧客業務の生産性低下
作業内容予測	技術伝承コストの削減	現場での保守業務にかかる時間を50%削減	保守作業者の減少と製品の複雑化に伴う技術伝承コストの増加

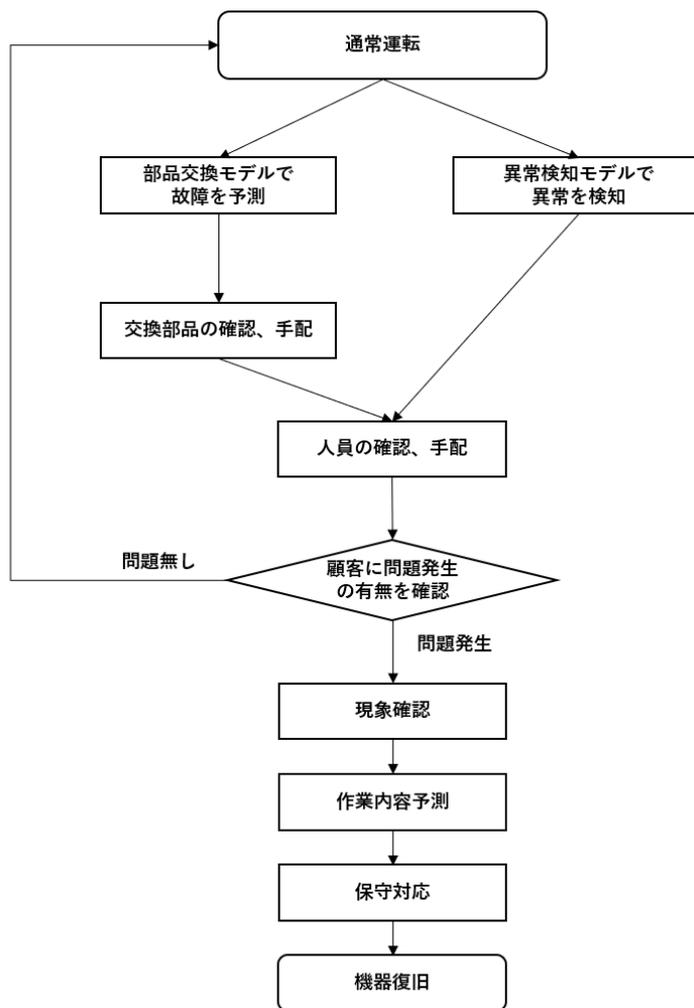


図2 構築する保守作業支援システム

### 3.1. 故障予測機能について

本研究で提供する故障予測機能では、異常検知モデルと部品交換予測モデルの2種類の故障予測モデルを作成することで、計画的な人員と交換部品の手配に活用していく。故障予測機能におけるシステムの構成を図3に示す。また、予測モデルの構築方法の詳細については4章で述べる。

部品交換予測モデルでは、主に保守作業における交換部品の調達を効率的に行うことが期待される。交換部品の在庫が不足してしまうと保守業務の遅延が生じ、顧客の満足度を下げる可能性がある。一方で、交換部品の在庫が過剰にあると、部品の管理や維持に無駄なコストをかけてしまう恐れがある。交換部品の需要を正確に把握することが出来れば、計画的な在庫管理と人員の手配を行うことが可能である。部品交換予測モデルの作成には、交換が必要な部品の種類を予測する必要があり、教師あり学習を用いて予測することが有効であると考えられる。しかし、教師あり学習を用いた先行研究では発生頻度の高い故障では予測に成功しているが、発生頻度の低い故障においては、データ不足等の要因により予測が行えていない問題点がある。そこで、最も交換頻度の高いインクヘッドを対象に部品交換予測モデルの構築を行う。

異常検知モデルでは、教師なし学習の異常検知を用いて機器の全般的な異常を検知することで、保守作業における人員の手配を行うことが期待される。教師なし異常検知とは、センサーの情報から正常なデータの特徴を学習し、正常なデータ分布からの外れ値を異常として検出する方法である。一般的に故障の記録は非常に少なく、教師あり学習を用いて全ての故障を予測することは困難であるが、教師なし異常検知では訓練データに正常なデータのみを使うため、故障記録の少ない異常を検知することが可能である[9]。保守業務に必要な人員の手配には、機器が異常か正常かを判別することが出来ればよいため、教師なし学習の異常検知が人員の手配に有効であると考えられる。他にも異常検知モデルでは、部品交換予測モデルで行う予測の信頼性を高める役割を期待して

いる。異常検知モデルと部品交換予測モデルでは、検出可能な問題発生範囲は異なっている（図4）。ただし、全ての検出可能な現象の範囲が異なっているわけではなく、2つの予測モデルが検出可能な範囲では、より信頼性の高い予測を行うことが出来ると考えられる。

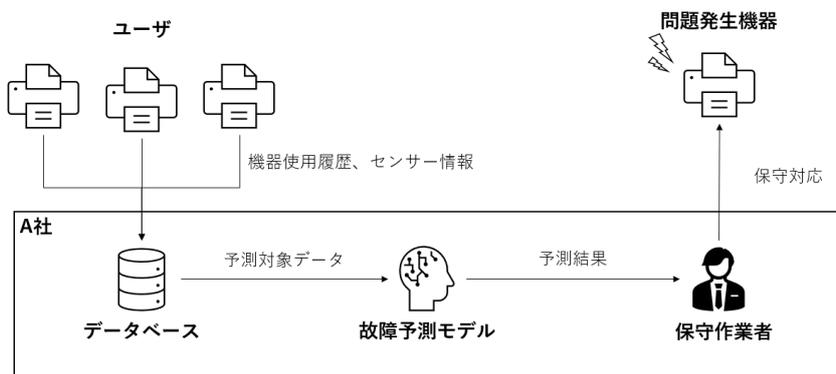


図-3 故障予測機能のシステム構成

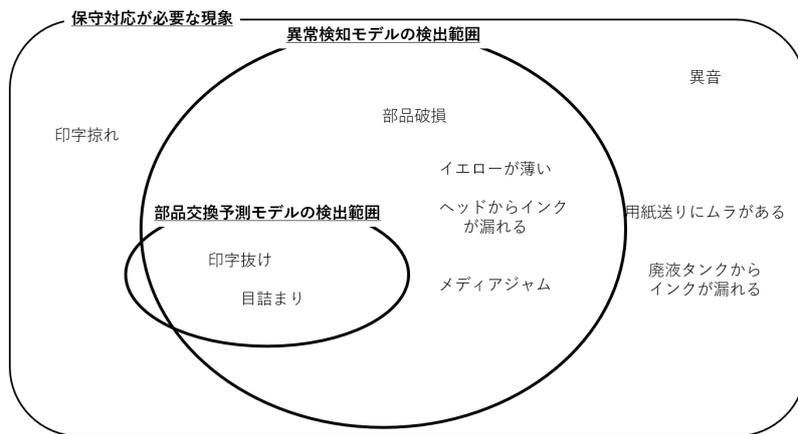


図-4 2種類の予測モデルによる問題発生の検出範囲例

### 3.2. 作業内容予測機能について

本研究で提供する作業内容予測機能では、保守作業記録を用いて作業内容の予測を行っていく。保守作業記録には問題発生時の現象、原因、対策が記載されている。本研究では、この保守作業記録が保守業務におけるノウハウであるとして、問題発生時の現象記録から作業内容を予測するモデルを構築する。このモデルを保守業務の際に使用することで、熟練作業員のノウハウを他の作業員でも活用できるように提供していく（図5）。これにより、技術伝承にかかるコストを削減し、問題発生の解決を効率的に行うことが期待される。



図-5 作業内容予測機能の構成

## 4. 部品交換予測モデルの構築

### 4.1. 対象データ

本章ではA社のプリンターインクヘッドの交換予測モデルを構築する。インクヘッドとは文字や写真などのデータをもとに、印刷対象にインクを吐出するプリンターのコア部品である。そのため、インクヘッドに問題が発

生した際には、印刷に支障をきたすため早急な対応が必要になる。また、インクヘッド交換は保守作業において最も交換回数が多く、最も高額な部品であるため、インクヘッド交換の予測がA社と顧客に与える影響は非常に大きい。

インクヘッド交換予測モデルの構築には、プリンター機器から不定期に送られてくるプロファイルデータを用いる(表 2)。プロファイルデータにはインク使用量やクリーニング状況、印刷枚数、カッター使用回数などの変数が存在する。データの取得期間はおよそ3~4年であり、この期間にヘッド交換が1,2回行われている。

また、本研究で使用するデータには予測を行う上で次の問題点が存在する。1点目はデータの取得間隔である。各プリンター機器から送られてくるデータの取得間隔が不定期であり、予測を行う際に補充が必要である。機体によっては1週間に1度しかデータが取得できていない機体が存在する。そのため、過度なデータ補間により予測精度が低下してしまうことが考えられる。2点目は予測に用いる変数が膨大なことである。プロファイルデータは125次元の変数が存在しており、故障の予測を行うには、重要な変数を選択することやデータを低次元に削減する処理が重要であることが考えられる。

表-2 データ概要

	次元数	データ件数	故障件数	機体数
前処理前	125	3748	18	39
前処理後	302	15877	17	14

#### 4.2. 前処理

##### (1) サンプル間隔の変更

本研究で扱うデータの取得間隔は不定期であり、機械学習を用いてデータを分析するには、サンプル間隔を一定にする必要がある。本研究では質的変数に対しては近傍補充を行い、量的変数に対しては線形補充を行い、サンプル間隔を1日毎に変更した。

##### (2) データ数の少ない機体の削除

取得したデータには、データ数が少ない機体で2件、多い機体で300件以上と、データ数に大きな差が見られた。データ数の少ない機体では、上記のサンプル間隔を変更する際のデータ補充が行えないと考え、補充前のデータ数が100件以下の機体は学習に使用せず、100件より多い機体を学習に使用した。

##### (3) カウンター系変数の初期化

大判インクジェットプリンタには印刷枚数や使用したインクの量、カッターの使用回数といったカウンター系変数が多く存在している。これらをヘッド交換毎に初期化することで、予測モデルにヘッド交換からの使用状況を適切に学習させることが期待される。

##### (4) 移動平均、階差系列

移動平均と階差系列をとることで、データの長期的な変動と短期的な変動を捉え、予測に有効な特徴を得られることが期待される。本研究では、移動平均と階差系列を各モータ寿命、カッターなどのカウンター系変数、各ヘッドのインク消費量に適用し、特徴量として加えた。

##### (5) 次元圧縮

本研究ではIsomapを用いて次元削減を行った。Isomapとは非線形の次元削減手法の一つでK近傍グラフを用いて多様体上の測地線距離を求め、多次元尺度構成法を使い近似的に低次元空間に射影を行うことで次元を圧縮する方法である。本研究で使用するデータは前処理後で300次元以上の高次元なものであるため、次元削減を用いて予測に有効な特徴を低次元の変数で表すことで精度の向上が期待される。

#### 4.3. 部品交換予測モデルの作成

各プリンター機器から送信されてくるプロファイルデータと機械学習を用いてインクヘッド交換の予測を行い、保守業務への有効性を検証する。保守において故障検出が早すぎる場合、機器が本来であれば稼働できたはずの期間を無駄にする恐れがある。そこで本研究では、ヘッド交換が起こる30日以内に交換を予測することを目標に行っていく。予測には教師あり学習のクラス分類を用いる。使用する教師データはヘッド交換日をクラス1、それ以外の日をクラス0として作成した。使用するデータの70%を訓練データ、30%をテストデータとする。訓練データには11件の交換記録とその間の正常データ、テストデータには6件の交換記録とその間の正常データが含まれている。次に実験の手順を示す。

##### (1) 特徴量選択と予測

データの次元数が膨大であるため、予測に重要な特徴量を選択する。特徴量選択には機械学習の中でもホワイトボックスな手法である Random Forest と Tab Net を使用する。この2つと前処理後の全ての変数を用いて交換予測を行い、各変数の重要度を求める。

(2) 特徴量選択後の予測

両手法における重要度の上位 10 個ずつ、最大で 20 個の特徴量を用いて Gated Recurrent Unit で最終的な交換の予測を行う。

4. 4. 実験結果

4. 4. 1. Random Forest による特徴量選択と予測

機体 B, C, E では交換確率がヘッド交換日に近づくにつれて、徐々に上昇する傾向がみられた(図 6)。また、テストデータにおいてヘッド交換日に交換確率が 50%を超えたのは 6 回中 3 回で、ヘッド交換日以外に交換確率が 50%を超えたのは 903 回であった (図 7)。変数の重要度については、ヘッド交換日毎に初期化したカウンター系変数の重要度が高い結果となった(表 3)。

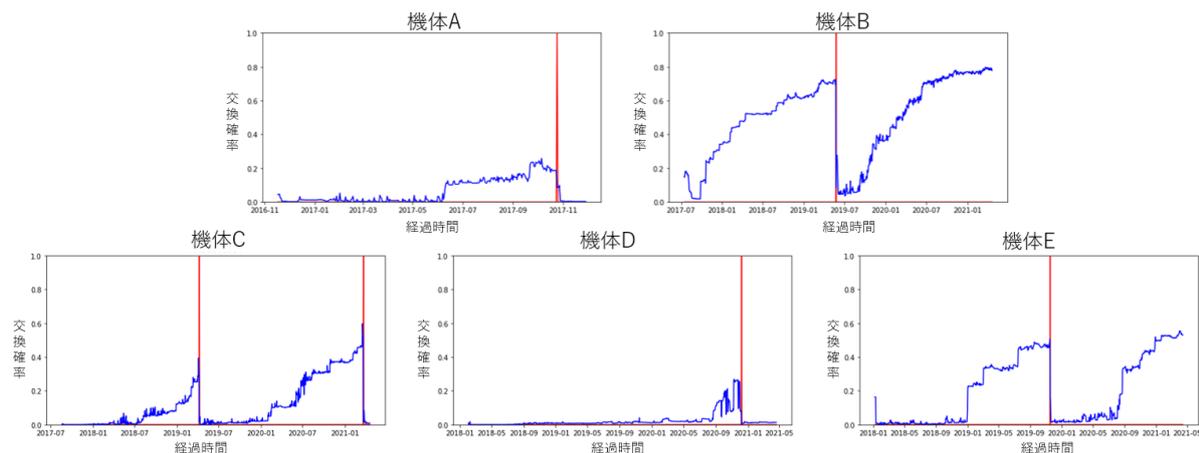


図- 6 Random Forest による予測結果、縦軸：故障確率、横軸：経過時間、赤線：ヘッド交換日（ヘッド交換の行われた日を 1、それ以外の日を 0）青線：予測モデルによって得られた故障確率

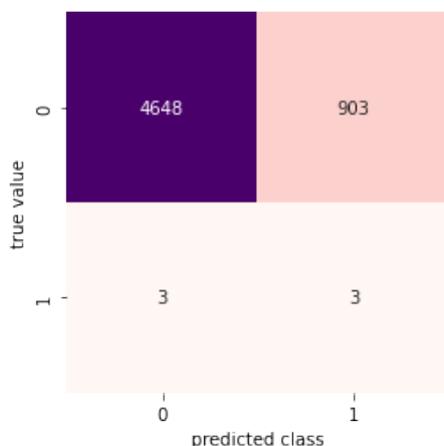


図- 7 Random Forest による全機体の予測結果（混同行列）

表- 3 Random Forest から算出された重要度上位 10

変数	重要度
データ長_cvlog_reset	0.0240
カラー番号5のインク消費量_cvlog_reset	0.0236
カラー番号4のインク消費量_cvlog_reset	0.0233
timestamp_cvlog_reset	0.0228
キャリッジY方向モーター_cvlog_reset	0.0179
ポンプモータAのステップ数_cvlog_reset	0.0171
累積印字タイマ_cvlog_reset	0.0170
原点側廃液のインク量	0.0169
カラー番号2のインク消費量_cvlog_reset	0.0166
カラー番号インク消費平均_cvlog_reset	0.0160

※cvlog\_reset : ヘッド交換日に初期化したカウンター系変数、mean:一週間ごとの移動平均

#### 4. 4. 2. Tab Net による特徴量選択と予測

ヘッド交換日に近づくにつれて交換確率が徐々に上昇する傾向はみられなかった (図 8)。また、テストデータにおいてヘッド交換日に交換確率が 50%を超えたのは 6 回中 2 回で、ヘッド交換日以外に交換確率が 50%を超えたのは 361 回であった (図 9)。変数の重要度では、Random Forest と同様に、ヘッド交換日毎に初期化した変数に加えて、一週間ごとの移動平均をとった変数の重要度が高かった (表 4)。

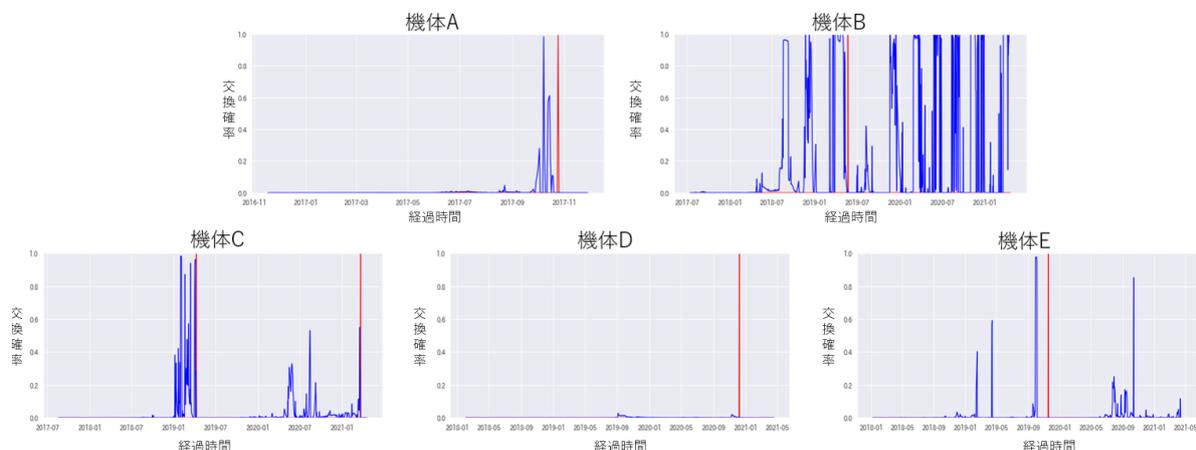


図- 8 Tab Net による予測結果、縦軸：故障確率、横軸：経過時間、赤線：ヘッド交換日（ヘッド交換の行われた日を1、それ以外の日を0）青線：予測モデルによって得られた故障確率



図- 9 Tab Net による全機体の予測結果（混同行列）

表- 4 Tab Net から算出された重要度上位 10

変数	重要度
カラー番号インク消費合計_cvlog_reset	0.2815
ヘッド 1 インク消費合計_mean	0.1721
ポンプモータ寿命：累積_cvlog_reset	0.1536
ポンプモータ寿命：累積_mean	0.1072
ヘッド1~4インク消費合計_mean	0.1049
ヘッド 2 インク消費合計_mean	0.0693
カラー番号 5 の充填されているインクの種類	0.0564
Isomap_neighbors15_1	0.0414
原点側廃液のインク量_mean	0.0043
カラー番号5のインク消費量_mean	0.0005

#### 4.4.3. Gated Recurrent Unit による予測

予測に用いる変数は Random Forest から得られた重要度の上位 10 個と Tab Net から得られた重要度の上位 10 個を用いて行った。機体 C, D, E では交換確率がヘッド交換日に近づくにつれて、徐々に上昇する傾向がみられた (図 10)。また、テストデータにおいてヘッド交換日に交換確率が 50%を超えたのは 6 回中 5 回で、ヘッド交換日以外に交換確率が 50%を超えたのは 1232 回であった (図 11)。Random Forest, Tab Net と比較して最も見逃しを少なくヘッドの交換を予測することが出来た。

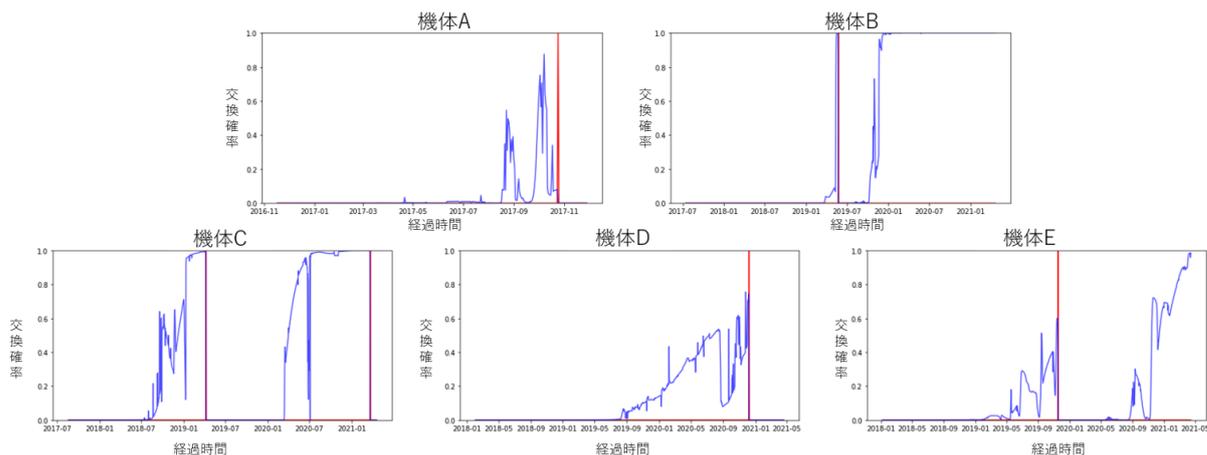


図- 10 Gated Recurrent Unit による予測結果、縦軸：故障確率、横軸：経過時間、赤線：ヘッド交換日（ヘッド交換の行われた日を1、それ以外の日を0）青線：予測モデルによって得られた故障確率

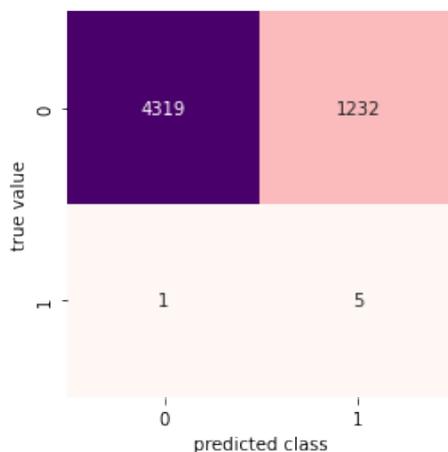


図- 11 Gated Recurrent Unit による全機体の予測結果（混同行列）

#### 4.5. 考察

今回の実験で行った交換予測では、ある日には交換確率が50%を超えていたが、次の日には交換確率が下がってしまうことがある。そのため、交換確率が50%を超えた時点からヘッド交換までの日数を、予測モデルによるヘッド交換検出期間とする(図12)。表5に各機械学習手法によるインクヘッド交換の検出期間を示す。

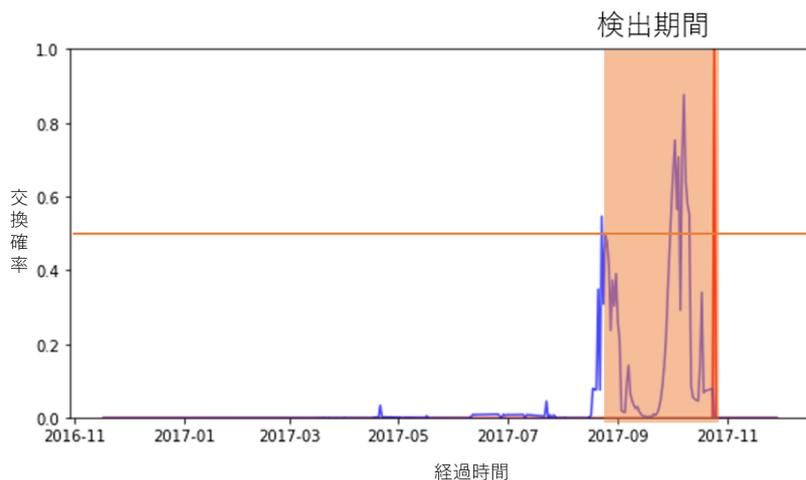


図- 12 故障検出期間の例

表- 5 各機械学習手法による検出期間の比較

	機体A 1回目	機体B 1回目	機体C 1回目	機体C 2回目	機体D 1回目	機体E 1回目
Random Forest	×	403	×	3	×	1
TabNet	215	×	310	372	×	180
Gated Recurrent Unit	23	11	201	354	135	5

重要度の高い特徴量について、ヘッド交換日毎にカウントをリセットした変数が Random Forest と Tab Net において重要度が高い結果となった。理由として次の2点が考えられる。1点目は、今回使用したデータに欠損値が多かったため、補間を行いやすいカウンター系変数の重要度が高いと判断されたこと。2点目は、カウンター系変数をヘッド交換日ごとに初期化したことで、インク消費量の値やモータ寿命値の組み合わせがインクヘッドの経年劣化を表せたと考えられる。

Gated Recurrent Unit では6回中6回の故障を検出することができ、他の機械学習手法と比較してインクヘッド交換日の見逃しが最も少ない結果となった。この理由として、Gated Recurrent Unit には時系列を考慮した学習が可能であるため、他の手法と比較して予測精度が高くなったことが考えられる。

本予測の目標はヘッド交換日の30日以内に故障を検知することである。Gated Recurrent Unit ではヘッド交換日の30日以内に6回中3回のヘッド交換を検出した。また、A社で働く社員に今回の予測に対する評価をいただいた結果、Gated Recurrent Unit を用いた故障予測はメンテナンスに有効であると判断された。また、検出期間があまりに長すぎるものは、異常検知モデルから算出された異常度との比較を行うことで予測に対する信頼性の向上を図っていく。

#### 4.6. 予測モデル改善の方策

本研究で使用したデータは、予測モデル構築のために取得したものではない。そのため、今後は下記の方策を行うことでシステム実装へ向けて取り組んでいく。

##### (1) データ取得間隔の改善

予測に使用したデータは各プリンター機器から不定期に取得しているため、機体によっては一か月に一度しかデータを取得していない機体もある。部品交換予測モデルの構築では、データ数が100件未満の機体を削除することや、線形補完や近傍補完を行うことで欠損値に対する処理を行っている。しかし、データの欠損が大きすぎる場合、期待通りの補完を行うことが出来ずに予測精度を低下させる原因となる。データの取得間隔をより短くすることができれば、データの欠損を最小限に抑え、補間による予測精度の低下を防ぐことが期待される。加えて、本実験ではカウンター系変数といった補完を行いやすい変数の重要度が高い結果であったが、データ取得間隔が短くなればヒータ温度等の短期的に変化する変数を扱うことが出来るようになるため、予測精度向上に繋が

ることが期待される。一方でデータ取得間隔を短くした際のデメリットとしては、より大量のデータを保存するためのサーバが必要になる点や大量のデータを送信するためのコストがかかってしまう点が考えられる。

(2) 特徴量選択手法の改善

本実験での特徴量選択はRandom Forest と Tab Net から得られる重要度をもとに行った。それぞれの手法で3回から4回のヘッド交換を予測できているが、全ての交換を予測できているわけではない。Random Forest と Tab Net の重要度のみを用いて特徴量選択をしてしまうと、本来重要であるはずの特徴量を逃してしまう恐れがあり、他の手法を用いて特徴量を精査していくことが重要である。一例として、熟練作業者に故障原因のヒアリングを行うことが挙げられる。熟練作業者は保守業務を通じて故障原因についての理解を深めている。そのため、熟練作業者が考える故障原因を予測の変数に加えることで予測精度向上に繋がることが考えられる。他にも、熟練作業者に本実験での特徴量選択で得られた特徴量についての意見を求め、予測に必要な特徴量を見極めることで、予測精度の向上に繋がることが期待される。

(3) 新規センサーの取り付け

機器の故障は主に初期故障期、偶発故障期、摩耗故障期の3つに区分することが出来る。それぞれの故障の主な原因として、初期故障期では製造時の欠陥や検品漏れ、偶発故障期では軽微な欠陥の悪化や人間のミス、摩耗故障期では機器の摩耗や疲労によるものが挙げられる。今回の予測ではカウンター系変数の重要度が高い結果となった。カウンター系変数は機器の使用経過を表すことができ、それら変数の組み合わせがヘッドの摩耗故障を予測したことが考えられる。そのため予測精度を上げるには摩耗故障以外の初期故障や偶発故障の予測が必要である。初期故障や偶発故障の原因は、製造時の欠陥や人間のミスであり、本実験のようにカウンター系変数が重要である可能性は低い。そこで、機器に振動や温度、音を計測するためのセンサーを取り付け、予測を行うことで初期故障や偶発故障を検知することが期待される。ただしセンサーを取り付けるためには、適切なセンサーの種類や取り付け位置を検証していく必要がある。加えて、既に販売してしまった機体にセンサーを取り付けることが難しいことやセンサー自体の費用がかさんでしまう問題点もある。

(4) 予測精度の調査

A社で働く社員の方に今回の予測に対する評価をいただいた結果、Gated Recurrent Unit を用いた故障予測は本研究で使用したデータに関しては、メンテナンスに有効であると判断された。しかし、本実験で構築した予測モデルを活用して保守業務を行うには、新たに取得したデータを用いて予測の精度を調べていく必要がある。

5. 作業内容予測モデルの構築

5.1. 対象データ

作業内容の予測には保守作業記録を用いて行う。データ概要は表6に示す。保守作業記録には問題発生時の現象、原因、対策、作業内容カテゴリが記録されている。本実験で使用する保守作業記録は72件のみで、予測対象の作業カテゴリの件数にはばらつきがある。そのため、本実験での予測対象は作業件数が、上位3つの作業カテゴリに限定する。予測を行う作業カテゴリ毎のデータ件数は表7に示す。

表-6 保守作業記録の概要

現象	原因	対策	作業カテゴリ
印刷時の擦れ	ヘッドの不具合	ヘッド交換	部品交換
メディアにインクが垂れる	ヘッド周りの汚れ	クリーニング	清掃
...	...	...	...

表-7 作業カテゴリ毎の発生件数

部品交換	設置	清掃
37	12	7

本実験の目的は、技術伝承にかかるコストの削減を行う予測モデルを構築することである。予測モデルの構築には自然言語処理モデルであるBERTを用いる。学習には説明変数として問題発生時の現象、目的変数を作業カテゴリとして教師あり学習を行う。また、使用するデータの70%を訓練データ、30%をテストデータとしてモデルの構築を行う。

5.2. 作業内容予測モデルの作成結果

表9にテストデータにおける予測結果を示す。作業カテゴリの予測を行った結果、テストデータにおける予測の正解率は82.3%であった。作業カテゴリにおいて清掃の必要があるものを部品交換と予測してしまう結果となった。

表- 8 学習用データによる予測結果

		予測結果		
		部品交換	設置	清掃
正解データ	部品交換	26	0	5
	設置	0	8	0
	清掃	0	0	0

表- 9 評価用データによる予測結果

		予測結果		
		部品交換	設置	清掃
正解データ	部品交換	11	0	0
	設置	1	3	0
	清掃	2	0	0

### 5.3. 考察

今回の実験では問題発生時の現象から作業カテゴリを予測し、評価用データにおける正解率が82.3%であった。また、清掃における予測では1件も正解することが出来なかった。清掃の予測が出来なかった原因として、学習に使用したデータに含まれる清掃の件数が最も少なかったことが考えられる。他にも、現象の記録は保守作業者が自然言語で記述したものであるため、記載内容には個人差がある。実験で使用したデータにも、現象を細かく記載しているものから、一つの単語で記載が終わっているものまである。解決策として、作業記録への記載をチェックシートにすることで、保守作業者の個人差に関係なく均一の作業記録を取得することが期待できる。他にも、作業記録をカテゴリ変数として扱うことができ、より少ないデータでも学習可能なアルゴリズムの適用が期待される。

本実験の目的は作業カテゴリの予測によって技術伝承にかかるコストの削減を行うことである。今回は作業カテゴリの予測を行ったが、作業カテゴリの予測だけでは、どの部品に対して作業を行うのかが不明であるため、保守業務の支援につながる可能性は低いことが考えられる。今後、保守業務の支援を行うには作業カテゴリを予測するだけではなく故障の原因や作業箇所の特定制を行う必要がある。

### 5.4. 予測モデル改善の方策

本研究で使用したデータは、予測モデル構築のために取得したものではない。そのため、今後は下記の方策を行うことでシステム実装へ向け取り組んでいく。

(1) データの新規取得、データ拡張

データ不足が原因でテストデータにおいて清掃の予測を正解することが出来なかった。今後も新規の保守作業記録を取得していくことで予測に十分なデータの確保を行っていく。データ拡張に関しては、保守作業記録に記載されている単語をもとに、似た意味の単語に置き換える処理を行い、学習用データを拡張することで、予測精度を向上させることが期待される。また、保守作業者が過去に行った保守業務の内容をヒアリングすることで、保守作業記録にはないノウハウを取得することが期待される。

(2) チェックシートの作成

保守作業記録の記載内容には一単語の記載で終わっているものから詳細を記載しているものまで多くの個人差がみられた。予測精度が低下する原因として、記載が少なすぎると必要な情報が抜けてしまい、記載が多すぎると不要な情報が含まれてしまうことが挙げられる。そのため、自然言語での記載をクラスタリング等からカテゴリ化し、それらを用いてチェックシートを作成することで作業内容への記載を均一化することが必要である。また、自然言語を用いた予測モデルの構築には、言語の多様性や文脈理解の複雑さから膨大

なデータが必要となる。作業記録をカテゴリ化することで予測モデルの構築に自然言語を扱う必要がなくなり、Random Forest や SVM のような手法を適用することで予測精度の向上につながる。一方、作業記録をチェックシート化することのデメリットとしては、確認項目が膨大になってしまうことで作業時間が延びる恐れがある点や確認項目への詳細を記載することが難しい点が挙げられる。対策として、項目を確認する際に大分類から小分類に移行できるようなシステム設計を行うことや、それぞれの項目に自由記述欄を設けることでチェックシートに不足があった際の対策を行っていく必要がある。

### (3) 故障原因、必要な対策の予測

本実験では作業カテゴリの予測を行ったが、保守業務を行うには、故障原因の特定と対策が重要である。そのため、本実験で実施した作業カテゴリの予測だけでは、業務を行うためには不十分である。上記(2)で記載した内容を実施することで故障原因や対策への記載をカテゴリ化し、故障原因や必要な対策の予測を行うことで保守業務に活用可能な予測モデルの構築を行っていく。理想的な予測モデルとしては、保守作業者がチェックシートへ発生している現象を記入する。その後、予測モデルから故障の原因と必要な対策が出力され、保守業務に取り組むことが出来るものである。そのため、今後は現象記録から原因と対策を予測するモデルの構築に取り組む。

## 6. まとめ

構築を目指す保守作業支援システムには2つの機能がある。1つ目は計画的なりソース確保の支援を行うための故障予測機能、2つ目は効率的にノウハウ共有を行うための作業内容予測機能である。本研究ではこのシステムの実現可能性の検証として、システムの中核となる2つの予測モデルの構築を行った。

部品交換予測モデルの構築では、Random Forest、Tab Net、Gated Recurrent Unit を用いて、インクヘッドの交換予測モデルを構築した。Random Forest、Tab Net から重要度の高い特徴量を選択し、Gated Recurrent Unit で予測することで、ヘッド交換が発生する30日以内に6回中3回の交換を検出することが出来た。また、Gated Recurrent Unit による予測をA社の職員に評価していただき、メンテナンスに有効であると判断された。特徴量選択ではカウンター系変数の重要度が高い結果となり、インクヘッドの摩耗故障の予測を行うことが出来たと考えられる。ただし、Random Forest と Tab Net の重要度のみを用いて特徴量選択をしてしまうと、本来重要であるはずの特徴量を逃してしまう恐れがあり、熟練作業者とのヒアリング等を用いて特徴量を精査していくことが重要である。予測精度が低下してしまった原因としてはデータの欠損が考えられる。データの欠損が大きすぎる場合、期待通りの補完を行うことが出来ずに予測精度を低下させる原因となる。データの取得間隔をより短くすることができれば、データの欠損を最小限に抑え、補間による予測精度の低下を防ぐことが期待される。一方でデータ取得間隔を短くした際のデメリットとしては、より大量のデータを保存するためのサーバが必要になる点や大量のデータを送信するためのコストがかかってしまう点が考えられる。

作業内容予測モデルの構築では、問題発生時の現象から作業カテゴリの予測を行い、正解率は82.3%であった。また、清掃に関する予測は1件も正解することが出来なかった。原因として、清掃に関する記録が7件しかなく、予測対象である作業カテゴリの中で最も少なかったことや、作業記録の記載に個人差が大きいことが挙げられる。作業記録への記載が少なすぎると必要な情報が抜けてしまい、記載が多すぎると不要な情報が含まれてしまうことで予測精度が低下する。また、自然言語を用いた予測モデルの構築には、言語の多様性や文脈理解の複雑さから膨大なデータが必要となる。そのため、自然言語での記載をクラスタリング等からカテゴリ化し、それらを用いてチェックシートを作成することで作業内容への記載を均一化することが期待される。加えて、作業記録をカテゴリ化することで予測モデルの構築に自然言語を扱う必要がなくなり、Random Forest や SVM のようなデータ数が少なくても予測可能な手法を適用することで精度の向上につながる。一方、作業記録をチェックシート化することのデメリットとしては、確認項目が膨大になってしまうことで作業時間が延びる恐れがある点や詳細を記載することが難しい点が挙げられる。対策として、項目を確認する際に大分類から小分類に移行できるようなシステム設計を行うことや、それぞれの項目に自由記述欄を設けることでチェックシートに不足があった際の対策を行っていく必要がある。本実験では作業カテゴリの予測を行ったが、保守業務を行うには作業カテゴリの予測だけでは不十分であるため、今後は現象記録から原因と対策を予測するモデルの構築に取り組んでいく。

今後は各予測モデル改善の方策を実施するとともに、異常検知モデルの構築や図2に示した保守作業支援システムの構築とその評価に取り組んでいく。

本研究の目的であるシステムの実現可能性の検証について、それぞれの予測モデルでデータ不足や欠損等の問題がみられたが、4.6節、5.4節で述べた予測モデル改善の方策を行っていくことで十分に実現可能であると考えられる。また、人材不足が引き起こす技術伝承コストの増加や長時間のダウンタイム発生等の課題を抱えている

企業は多い。本研究ではA社での課題を例に部品交換予測モデルの構築と作業内容予測モデルの構築について取り組んでいるが、教師あり学習と教師なし学習を使い分けて予測精度の向上を図る点や保守作業支援システムのフレームワークについては、同様の課題を抱えている企業に対しても有用であり、本研究で目的としたシステムの実現可能性の検証は、これら企業に対しても同様に有用であると考えられる。

[参考文献]

- [1] 経済産業省, “製造基盤白書 (ものづくり白書)”, 2018
- [2] 今村 誠, “予知保全のための機械学習”, システム/制御/情報, 65 巻, 4 号, p. 119-125, 2021
- [3] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton: Deep learning. Nature 521.7553 2015: 436-444
- [4] 松尾豊, 人工知能は人間を超えるか: ディープラーニングの先にあるもの: KADOKAWA, 2015.
- [5] Google: Cloud AutoML, <https://cloud.google.com/automl>
- [6] 梅津里香, 杉江卓哉, 長瀬雅之, et al, ” 機械学習を用いた宇宙機の故障の予兆検知、宇宙科学情報解析論文誌”, JAXA-RR-18-008, 8 号, p. 11 - 20, 2019
- [7] 切通 恵介, 泉谷 知範, “機械学習を用いた工場機器の故障予測”, DEIM forum 2017
- [8] 奥 武憲, 水野 哲志, 林 俊輔, 荒井 仁, et al. MFP のビッグデータを活用した故障予測”, Ricoh technical report, 43 号, p. 56-62, 2018
- [9] 曾我部東馬: Python による異常検知, オーム社, 2021

(2023 年 8 月 18 日受理)