

プロセス系製造に向けた不良原因発見のための データ分析手順の定型化

小松田 卓也 朝家 真知子

(株) 日立製作所 研究開発グループ

1. はじめに

センシング技術やコンピューティング性能の向上により、製品、環境、作業等に関する様々なデータを収集・分析できるようになってきており、売上向上や仕損費削減等の経営効果を得ている事例が増えてきている[1]。例えば、製造業向けに、製品規格、装置稼働ログ、製品検査に関するデータを分析し、不良の原因を特定して、歩留りを低減する取組みが行われている。

従来、分析案件毎に要件や分析担当者が異なることが多く、分析手法をその都度カスタマイズするケースが多かった。しかし、これまで担当してきた複数の分析案件を振り返ってみると、分析手法に共通部分が存在することが分かってきた。上記のような背景から、我々は、下記手順に従い、複数の分析案件に適用可能な分析フロー及びツールの提供をめざしている。

- ① 過去類似案件間で共通の分析ノウハウを抽出し、データ分析手順を定型化する。
- ② ①で定型化した分析手順に関して、他の案件に適用し、所望の効果を得られるか検証する。
- ③ 案件共通での利用を可能にするため、データ分析手順の適用範囲を拡大する。

本研究では、まず①に取り組んだ。具体的には、データ利活用に積極的なプロセス系製造業を1stステップのターゲットとして、プロセス系製造向けのデータ分析手順を定型化した。分析手順を実問題に適用した結果、分析手順の設計における試行錯誤が不要になったため、37.5%の工数を削減できることを確認した。以降では、2章でプロセス系製造におけるデータ分析、3章で提案、4章で検証、5章でまとめを述べる。

2. プロセス系製造におけるデータ分析

製造業において、歩留り向上に向けた未知の不良原因発見や稼働率向上に向けた設備故障の予兆検知が求められている。これら課題の解決に向け、データ利活用が注目されており、多くの工場では設備にセンサを設置しデータを収集し始めている[2]。特にプロセス系製造業では、製造工程において、人手の介在余地が少なくほとんど設備が担っているため、設備稼働データから価値抽出できることを期待しており、設備データを積極的に収集している。設備データの利活用に向けて、所望の結果を得るために、データ加工や統計手法の選

択等を正しい順序で実施する必要がある。CRISP-DM(Cross-Industry Standard Process for Data Mining)[3]等の標準的なデータ分析方法論に従うことが一般的である。しかし、CRISP-DMの実案件適用にあたって、分析案件毎に収集データのセキュリティやデータ管理ポリシーが異なり、それに伴って分析方法も異なるため、分析案件個別に分析手順の設計に試行錯誤が必要となり、1件の分析案件あたりの工数が高くなる。本研究では、分析手順の設計工数を削減するために、分析手順の定型化を行い、データ分析工数の削減に取り組む。

3. 提案

製品の不良原因特定に向けた製造条件判別ルール作成のためのデータ分析に関して、分析案件共通の分析手順の抽出を行い、各手順における作業チェックシートを作成した。図1に分析案件共通の分析手順を示す。

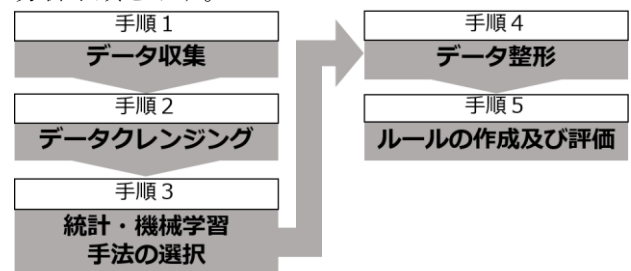


図1：データ分析手順

データ収集、データ理解、統計・機械学習手法の選択、データクレンジング、ルール作成及び評価の順に実施する。以下に各作業の概要を示す。

手順1：データ収集

まず、製造データ（製品製造時におけるセンサ値、使用設備名等のデータ）と品質管理データ（製品の良・不良判定結果）を収集し、上記2種類のデータを紐付けることで、不良品製造時における設備情報（温度、ライン速度等）を作成する。そのために、下記3点を確認する。

- ・ 製造データを収集できているか
- ・ 品質管理データを収集できているか
- ・ 製造データと品質管理データの紐付けができるか

上記項目の未達成のものに関しては、収集を行った上で次ステップに進む。

手順2：データクレンジング

センサ故障やデータ入力者の誤記入により、正しくデータを収集できておらず、重複値や外れ値、欠損値を含む場合がある。まず、それらデ

Proposal of Data Analysis Template for Defect Cause Detection in Process Manufacturing
†Takuya Komatsuda, Machiko Asaie, Hitachi, Ltd. Research & Development Group

ータの有無について調査する。その後、現場データ管理者に重複値や外れ値の修正方法についてヒアリングし、代替値（直前の値や前後の値の平均値など）の入力や削除をする。

手順3：統計・機械学習手法の選択

不良品の製造条件判別ルールを作成するために、統計手法や機械学習手法を選択する。手法選択に当たって、数学的な特性上、手法毎にデータ特徴に制約が設けられており、製造データの分布や相関関係等を理解した上で選択する必要がある。データ特徴に関して、あらかじめ作業チェックシートに列挙し、例えば、センサデータは正規分布に従うか、センサ数が多いか等を確認する。

手順4：データ整形

手順3で選択した統計・機械学習手法について、製造データを手法の制約に合わせるために、データを整形する。例えば、ロジスティック回帰について、説明変数が正規分布に従うという仮定に基づいてルール作成するために、正規分布に従わないセンサを除去する。

手順5：ルールの作成及び評価

手順3で選択した手法を製造データに適用し、製造条件判別ルールを作成する。作成後、未知データにルールを適用し精度評価する。精度目標に達していない場合、センサデータの選定や手法のパラメータチューニング、期間による製造データの絞り込み等を行い、再度ルールを作成する。必要に応じて手順2, 3, 4に戻る。

次に、作業チェックシートに関して述べる。表1にデータ理解におけるチェック項目を示す。

表1：作業チェックシートの抜粋

大項目	中項目	#	詳細
データ理解	一貫性	2-1	命名規則に一貫性があるか確認 (一貫性のない例としては、あるカラムでは“ID”という表記である一方、別のカラムでは“Id”と表記されているなど)
		2-2	データ値に矛盾がないか確認 (開始時刻が終了時刻よりも遅いなど)
	データ型	2-3	カラム毎のデータ型(数値、カテゴリID、テキスト)確認
		2-4	一つのカラムに複数のデータ型が混在していないか確認
	統計量	2-5	レコードのカウント
		2-6	カラムのカウント
		2-7	カラムに格納されている値の種類数のカウント
		2-8	ファイルの種類数のカウント
		2-9	外れ値があるか確認
	重複	2-10	同一内容ファイルがあるか確認
		2-11	同一内容レコードがあるか確認
		2-12	キーが重複しているか確認
	欠損	2-13	欠損ファイルがあるか確認 (IDや日付が飛んでいないかなど)
2-14		カラム毎に欠損値の確認	
2-15		カラム毎に欠損値の割合確認	
2-16		収集できていない区間(期間)を確認	

データ分析者は、例えば、データ理解に向けて、命名規則の一貫性があるかどうか（表1の2-1）やデータ値に矛盾がないか（表1の2-2）をシートに従って確認する。上記分析手順及びチェックシートを用いることで、

データ分析に関するドメイン知識なしに分析をすることができる。

4. 検証

プロセス製造業における不良原因特定に向けたデータ分析案件を対象に、提案分析手順を適用した。下記に適用結果を手順毎に述べる。

手順1：1年間分の製造データと品質管理データを使用した。

手順2：製品を製造していない時間帯のセンサ値の除外や、紐付けのための名前修正を行った。

手順3：100種類以上のセンサに対して、1センサずつルール作成すると工数を要する。分析対象を絞り込むために、主成分分析を選択した。また、良品及び不良品製造時にそれぞれ収集した製造データ（以降、それぞれ良品データ、不良品データと呼ぶ）について、ヒストグラムを比較して、センサ値の特徴の違いを抽出した。

手順4：主成分分析では、入力値として量的データを想定しており、質的データをそのまま適用することができない。そのため、装置の開閉フラグ等の質的データについては除外した。

手順5：センサ毎に良・不良データのヒストグラムをそれぞれ作成し、ヒストグラムの最頻値を比較した。比較した結果、最頻値の差が大きいセンサについては、不良原因に関係のあるセンサとして、工場現場作業者に報告することにした。

提案分析手順に従った結果、従来15人日要していたところ、9.375人日に削減できた（削減率37.5%）。

上記分析を実施した結果、設備の設定値を通常よりも低い値に変更した際に不良の発生率が高いという結果を得た。設定値を通常時の値に設定し直すことで不良発生を低減することが期待できる。

5. まとめ

複数業種に適用可能な分析フロー及びツールの提供に向けて、過去類似案件間で共通の分析ノウハウを抽出し、データ分析手順を定型化した。プロセス製造における不良原因特定に向けた実案件に提案手順を適用した結果、分析手順の設計における試行錯誤が不要になったため、従来に比べて工数を37.5%削減できた。今後、分析手順の適用範囲拡大に向けて、他の分析案件に適用可能か検証する。

参考文献

[1] <http://www.hitachi.co.jp/products/it/bigdata/case/>, 2017年12月25日アクセス
 [2] 2017年版ものづくり白書, <http://www.meti.go.jp/report/whitepaper/mo-no/2017/>, 2017年12月25日アクセス
 [3] Chapman, Pete, et al. "CRISP-DM 1.0 Step-by-step data mining guide." (2000).