

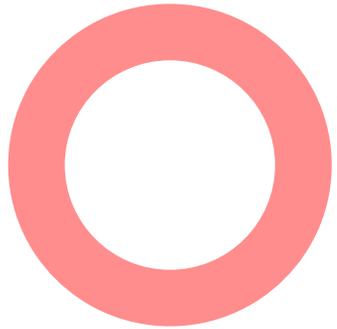


Sep. 28
2020

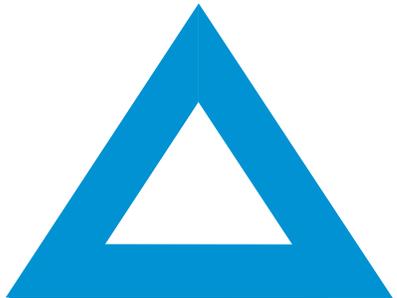
化学プラントにおけるソフトセンサー実装事例

ダイキン工業 TIC データ活用推進G 兼 化学事業部 プロセス技術部
伊與田 淳平

本発表でお役に立てるのは、以下の様なニーズをお持ちの方です



- ・データ解析の具体的な手順が分からない
- ・MATLABを使用したソフトセンサー構築の具体的な手順が知りたい
- ・MATLAB導入によって、どういった効果があるのか？



- ・新しい着眼点を持った特徴量抽出法が知りたい
- ・斬新なアルゴリズムを用いたソフトセンサーを見たい
- ・かっこいい画面のソフトセンサー構築方法ってどうやればいい？

1. 会社紹介
2. 背景情報/技術紹介
3. ソフトセンサー実装事例
4. 学んだこと/所感

1. 会社紹介
2. 背景情報/技術紹介
3. ソフトセンサー実装事例
4. 学んだこと/所感

(2020年3月末現在)

会社名	ダイキン工業株式会社 1963年（昭和38年）大阪金属工業株式会社から社名変更
創業	1924年（大正13年）10月25日大阪市で創業 創業者：山田晁
設立	1934年（昭和9年）2月11日
資本金	850億円
グループ従業員数	連結80,369名
会長・社長	会長：井上礼之 社長兼CEO：十河政則
本社	大阪市北区
グループ会社数	連結子会社313社（国内29社、海外284社）

創業**1924**年
96年の歴史

人を基軸に
おく経営

空調機器と冷媒を
両方手がけている
総合空調メーカー

グローバル生産拠点
100カ所以上

全社売上高
2兆円以上

従業員**8**万人
海外従業員比率は約**8割**

150カ国以上へ
事業展開

海外売上高比率
77%



空調事業



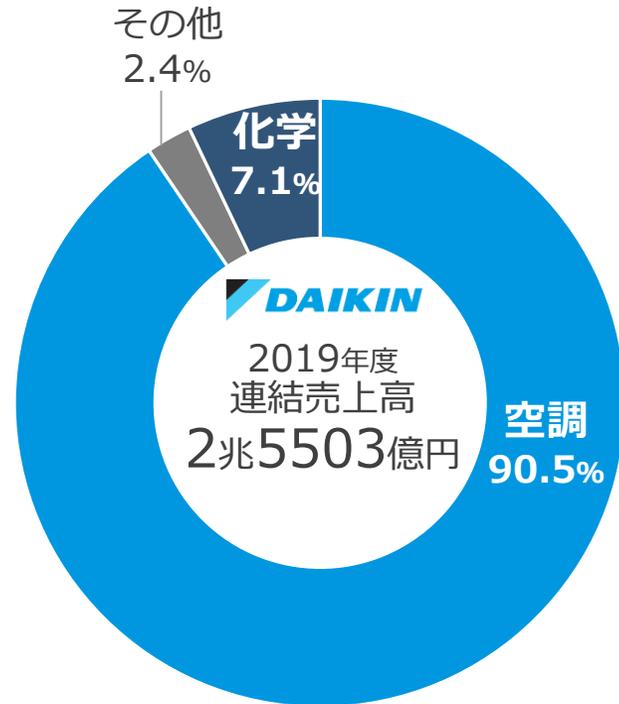
住宅用



業務用



サービス



その他事業



油圧機器

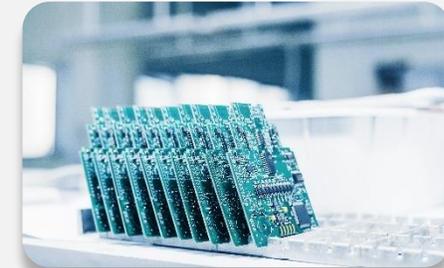


酸素濃縮機

化学事業



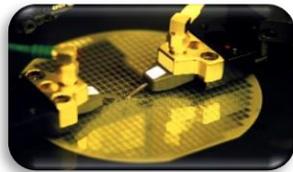
冷媒



半導体用途



自動車用途



半導体分野



ウエルキャリア
(PFA)



薬液ボトル/タンク
(PFA)



配管パイプ継手
(PFA・PTFE)



自動車分野



燃料ホース
(フッ素ゴム)



クランクシャフトシール
(フッ素ゴム)



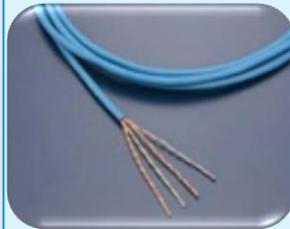
燃料噴射装置
(フッ素ゴム)



情報通信分野



タッチパネル
防汚(フッ素コート)



LANケーブル
(FEP)



プリント基板
(フッ素エポキシ)



環境エネルギー分野



リチウムイオン電池
(フッ素樹脂)



太陽電池
(ETFEフィルム)



風力発電
(ゼッフル)



生活・住宅分野



フライパン
(フッ素塗料)



炊飯器
(フッ素塗料)



ハンバーガー包装材
(紙用耐油剤)



住宅屋根・外壁
(ゼッフル)



エアコン
(冷媒ガス)



アパレル
(撥水撥油剤)

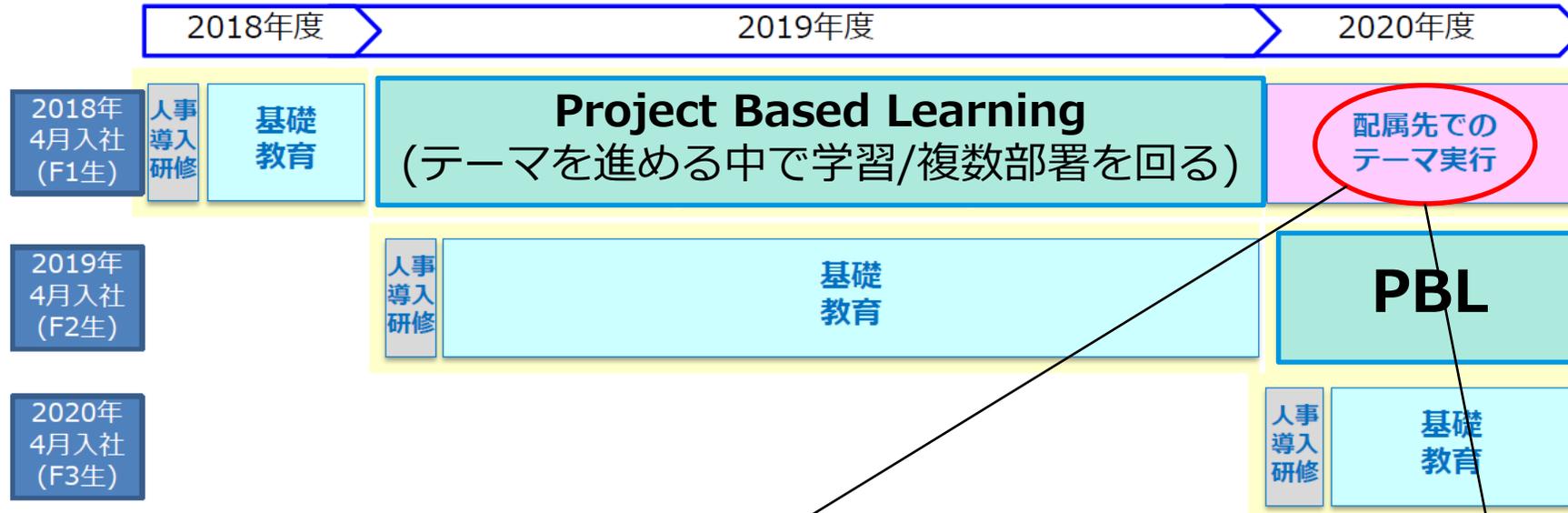
「タダ飯 ダイキン」

検索

2018年度から、通常の新卒採用にプラスして
毎年100人ずつ多く採用し、AI技術に関する社内研修を
2年間にわたって実施する、弊社内での人材育成戦略

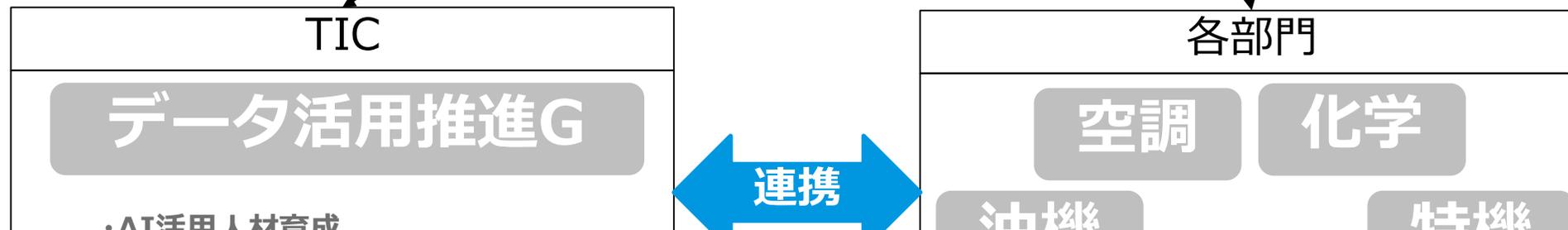
※週刊ダイヤモンドオンライン(2018/5/8)の記事

「タダ飯」スケジュール



※別途、2017年度より、既存社員のAI/IoT研修講座も実施

配属のイメージ(2020年度4月～)



データ/デジタル技術に強い人材を全社的に配置し、成果創出を狙う

※TIC：全社横断的な研究開発拠点（設立：2015/11）

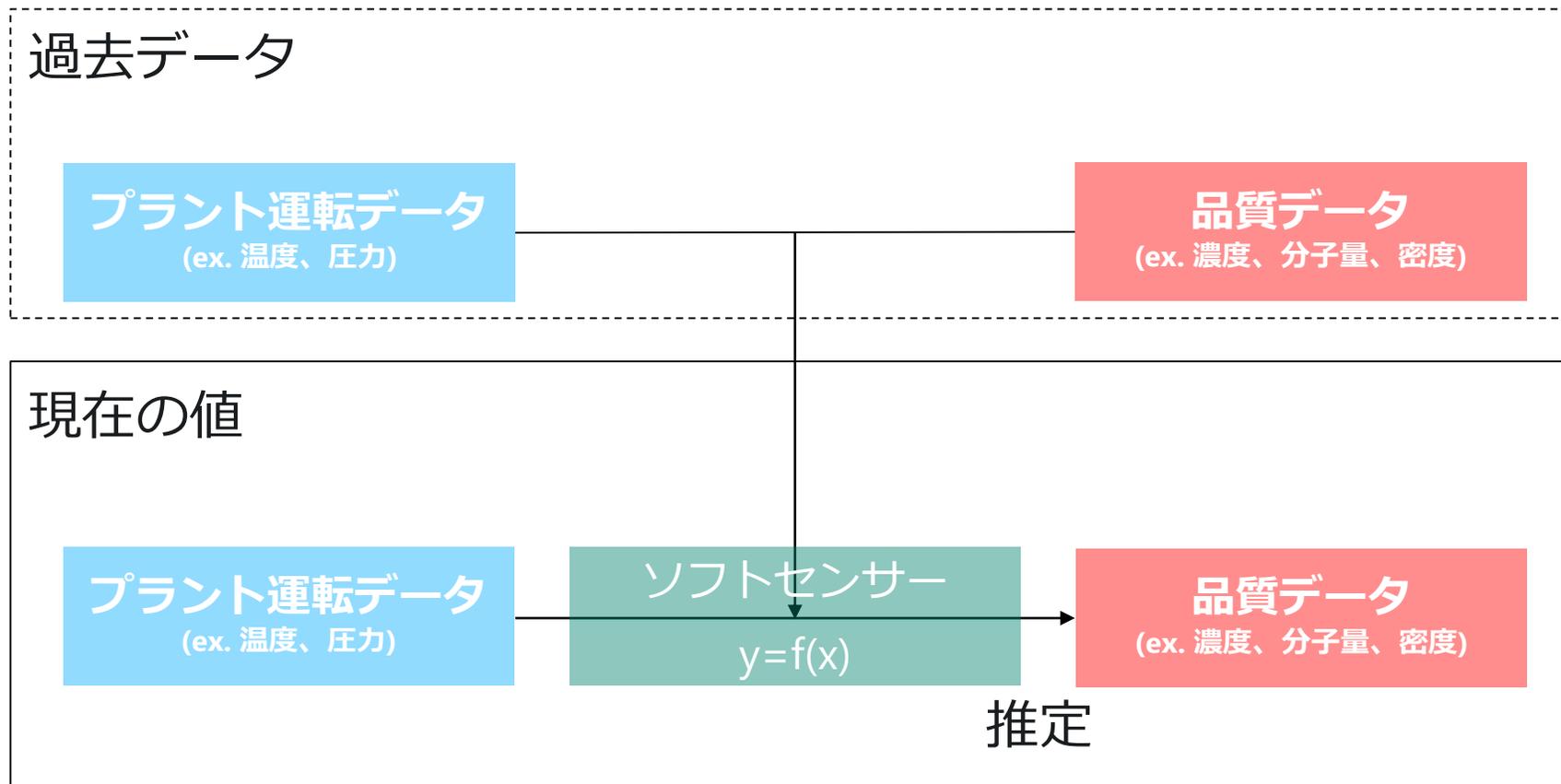
1. 会社紹介
2. 背景情報/技術紹介
3. ソフトセンサー実装事例
4. 学んだこと/所感



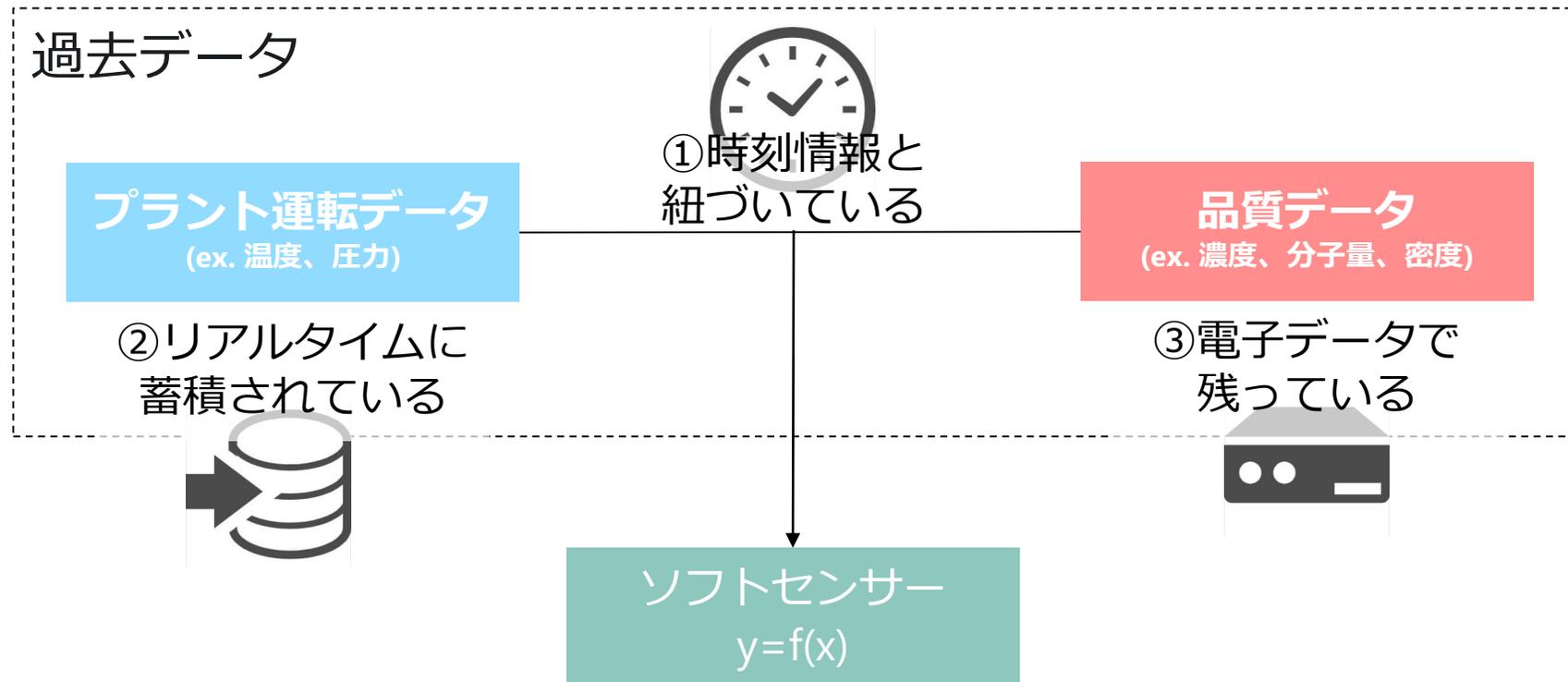
中身の品質が分かれば
最適なタイミングで
手直しができるのに・・・

- プロセス産業では、入口から出口まで閉じた系内で製造を実施。
不安安全
- 製造途中での品質確認のため、系内から製造中の製品を**取り出す**必要がある。
工数大
- 取り出した後、**工数をかけて手作業**で品質を測定しているケース多数。
- 品質の値によって**手直し**が**機会損失**

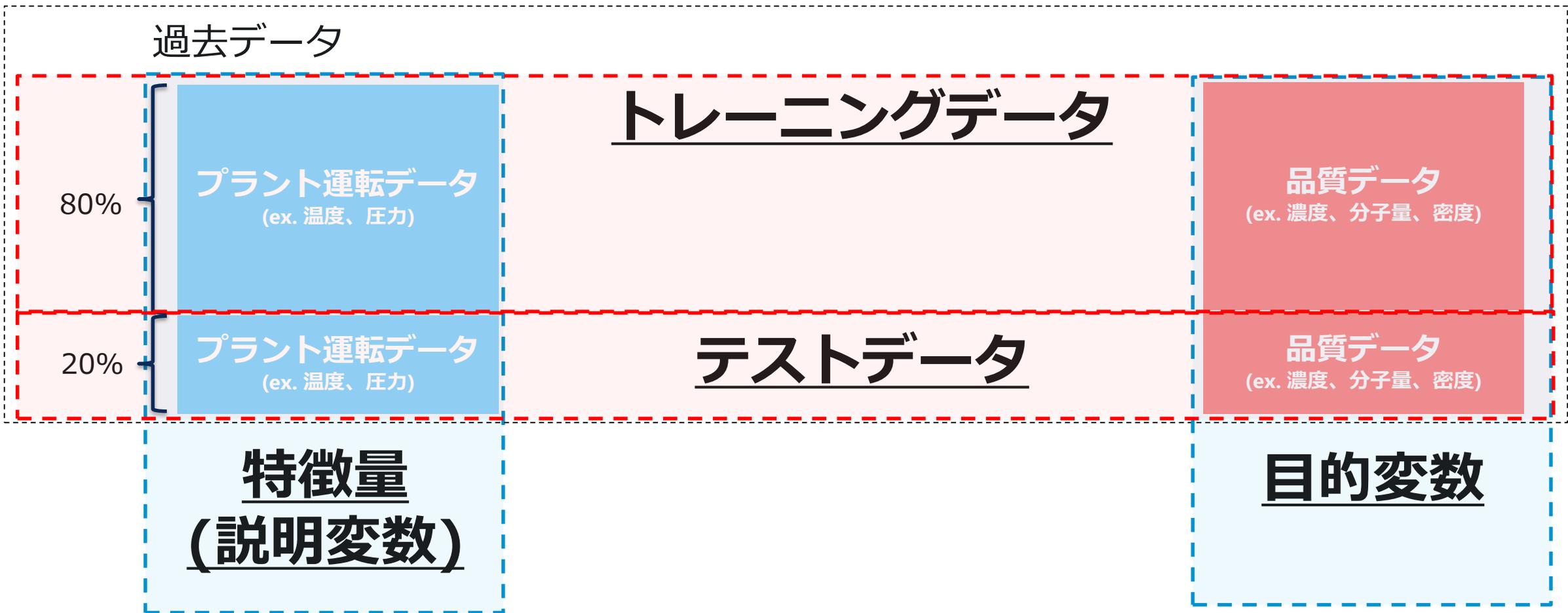
ソフトセンサー技術により上記課題を解決可能

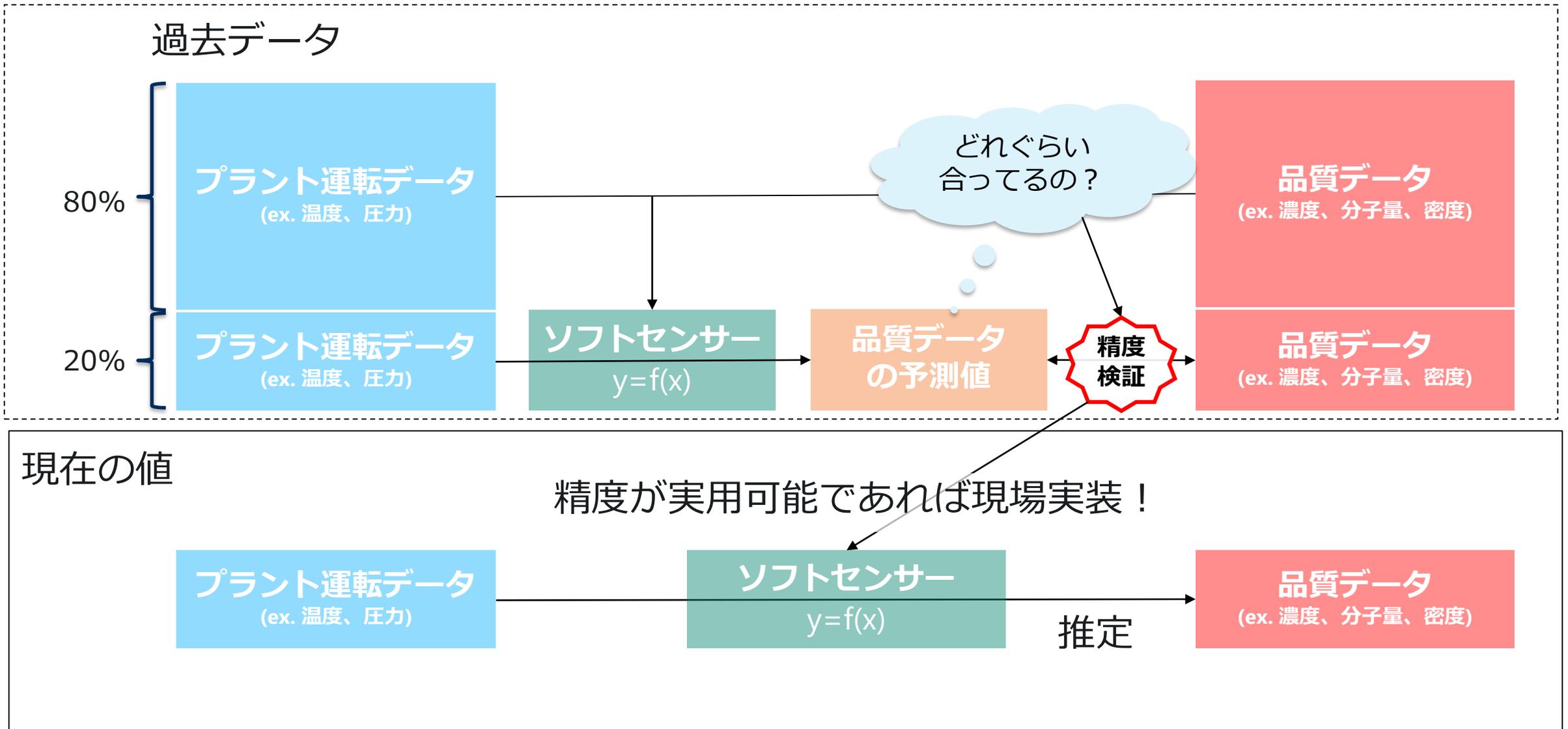


ソフトセンサーとは・・・
測定が難しいデータをリアルタイムに推定する「関数」のようなもの



上記3点を満たす事で、ソフトセンサー構築検討が開始できる





1. 会社紹介
2. 背景情報/技術紹介
3. ソフトセンサー実装事例
4. 学んだこと/所感

ヒアリング



- ・ 解決したい 課題を聞く
- ・ データの有無/
形/場所 など

データ収集



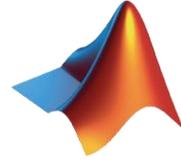
- ・ データ収集
- ・ 欠損データ確認

特徴量抽出



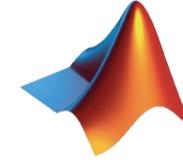
- ・ データ対応付け
- ・ データ加工
- ・ 予測モデル作成
できる形に
データ成形

予測モデル作成



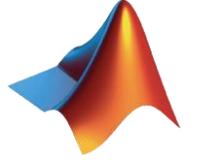
- ・ 最も 予測精度が
良い手法を
探索する

予測モデルの リアルタイム化



- ・ データベースと接続
- ・ リアルタイムに
予測値をプロットする

実装(アプリ化)



- ・ MATLABライセンス
が無い環境でも動く
アプリを作成

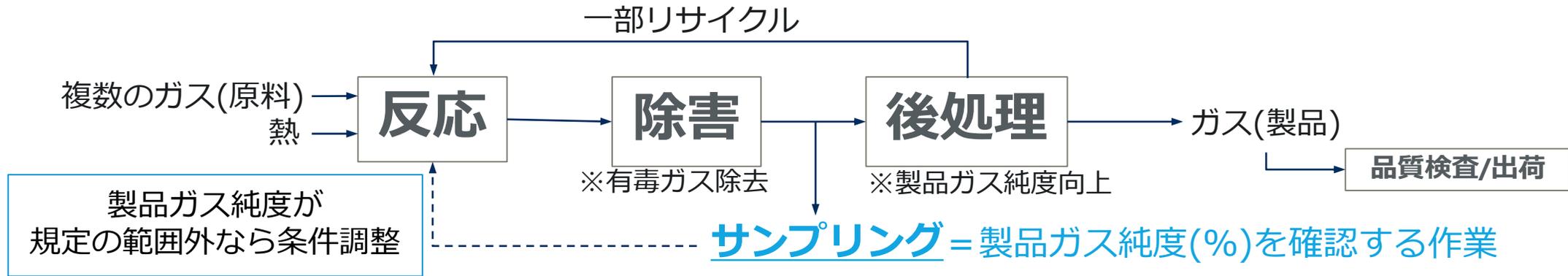
◆PI System

OSIsoft社が開発したPIMS（操業情報管理システム）
全世界のプラント産業で使われている。

弊社では主に
プラント運転データのデータベースとして使用。

上記手順を実施することで
プロトタイプ(試作モデル)作成
まで到達できる

◆本発表で取り扱うガス製造プロセス



◆サンプリング作業 (1日15分 × 1年間 × 20件以上)

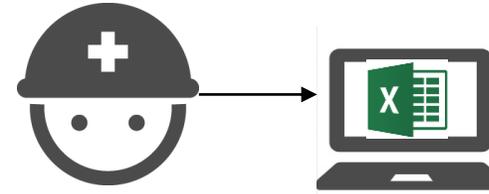
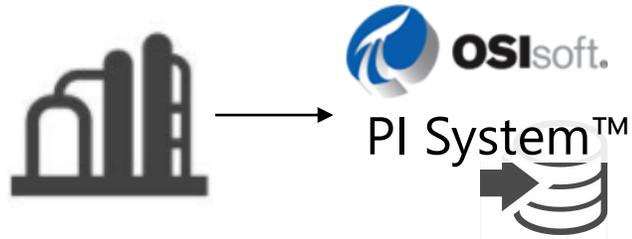


作業 工数削減/安全性向上のため
手作業でのサンプリングを無くしたい

◆現場の方からのヒアリング内容

- ・現場で、純度をコントロールするために使っているプロセスデータを蓄積している
⇒ 1分間隔のガス流量、反応温度データが時刻と紐づいて存在
- ・純度のデータもエクセルデータとして蓄積している
⇒ 1日ごとに手入力のエクセルデータが時刻と紐づいて存在
- ・製造工程の途中での傾向監視目的の検査のため、品質保証の意味合いは無い
⇒ プロセスにフィードバックするための工程内検査である
- ・ソフトセンサーとしての具体的な精度目標がある
(例：予測値と実測値の差が最大で2以下)

上記ヒアリング内容から、本工程は、
ソフトセンサー導入検討をする意義があると判断



◆ プラント運転データのデータベース
⇒ 温度、圧力、ガス流量、バルブの開閉 等のデータを1分間隔で自動で蓄積

◆ 純度データの蓄積
⇒ サンプルングして測定したガス純度データを約1日間隔で手動で蓄積

	A	B	C	D	E	F
1	time	撹拌電流値	ガスA流量	積算流量	ガスB流量	バルブA開閉
2	2018/12/27 8:25	0.160914421	-0.065386526	278402.4	91.10560608	CLOSE
3	2018/12/27 8:26	0.160860404	-0.065405443	278402.4	91.10560608	CLOSE
4	2018/12/27 8:27	0.160806388	-0.06542436	278402.4	91.10560608	CLOSE
5	2018/12/27 8:28	0.160752371	-0.065443277	278402.4	91.10560608	CLOSE
6	2018/12/27 8:29	0.160698354	-0.065462194	278402.4	91.10560608	CLOSE
7	2018/12/27 8:30	0.160644338	-0.065481119	278402.4	91.10560608	CLOSE
8	2018/12/27 8:31	0.160590321	-0.065500036	278402.4	91.10560608	CLOSE
9	2018/12/27 8:32	0.160536304	-0.065518953	278402.4	91.10560608	CLOSE
10	2018/12/27 8:33	0.160482287	-0.065537870	278402.4	91.10560608	CLOSE
11	2018/12/27 8:34	0.160428270	-0.065556787	278402.4	91.10560608	CLOSE
12	2018/12/27 8:35	0.160374253	-0.065575704	278402.4	91.10560608	CLOSE
13	2018/12/27 8:36	0.160320237	-0.065594621	278402.4	91.10560608	CLOSE

「反応」工程のみ

&

	A	B
1		純度
2	2019/5/2 8:58	85
3	2019/5/3 10:38	82
4	2019/5/4 10:05	82
5	2019/5/5 12:51	83
6	2019/5/6 13:02	86
7	2019/5/7 9:05	83
8	2019/5/8 9:45	84
9	2019/5/9 9:07	83
10	2019/5/10 9:21	83
11	2019/5/11 9:16	85
12	2019/5/18 11:08	84
13	2019/5/19 8:53	81
14	2019/5/20 8:55	85
15	2019/5/21 9:25	83

特徴量 ≡ 純度を予測するのに使えそうなデータ

	純度	特徴量1	特徴量2	特徴量3	特徴量4
2019/5/2 8:58	85	○○	××	△△	◆◆
2019/5/3 10:38	85	◎◎	×××	▽▽	□□
2019/5/4 10:05	82	●●	××××	▲▲	◇◇
2019/5/5 12:51	81				
2019/5/6 13:02	84				
2019/5/7 9:05	83				
2019/5/8 9:45	83				
2019/5/9 9:07	83				
2019/5/10 9:21	84				

...

作成

予測モデル
= ソフトセンサー

$$f(x)$$

ソフトセンサー作成のためには、
目的変数 1つに対して 1行の特徴量を用意する必要がある

純度データ ⇒ 約 1日毎

プラント運転データ ⇒ 1分毎

	純度
2019/5/2 8:58	85
2019/5/3 10:38	85
2019/5/4 10:05	82
2019/5/5 12:51	81
2019/5/6 13:02	84
2019/5/7 9:05	83
2019/5/8 9:45	83
2019/5/9 9:07	83
2019/5/10 9:21	84
2019/5/11 9:16	86
2019/5/18 11:08	83
2019/5/19 8:53	84



どの様に
関連付けるのが
適切か？

「関連付ける」作業
≡ 特徴量抽出

	Z1-CC0000.PV	Z1-CC0015.PV	Z1-CC0004.PV
	原料A 仕込み流量	反応器内 温度8	反応器 出口圧力
2019/5/2 8:58	18.3811	92.0571	0.6876
2019/5/2 8:59	18.3814	92.0582	0.6876
2019/5/2 9:00	18.3817	92.0593	0.6876
2019/5/2 9:01	18.3819	92.0604	0.6876
2019/5/2 9:02	18.3822	92.0615	0.6876
2019/5/2 9:03	18.3825	92.0626	0.6876
2019/5/2 9:04	18.3827	92.0638	0.6876
2019/5/2 9:05	18.3830	92.0649	0.6877
2019/5/2 9:06	18.3833	92.0660	0.6877
2019/5/2 9:07	18.3835	92.0671	0.6877
2019/5/2 9:08	18.3838	92.0682	0.6877

- ①純度データを記録した時刻のプラント運転データを使用
- ②純度データ記録時刻からさかのぼって
プラント運転データを取得。演算して特徴量に落とし込む

左記、2つの例を用いて
特徴量抽出について説明

①品質データを記録した時刻のプロセスデータを使用 のケース

純度データ

	純度
2019/5/2 8:58	85
2019/5/3 10:38	85
2019/5/4 10:05	82
2019/5/5 12:51	81
2019/5/6 13:02	84
2019/5/7 9:05	83
2019/5/8 9:45	83
2019/5/9 9:07	83
2019/5/10 9:21	84
2019/5/11 9:16	86
2019/5/18 11:08	83
2019/5/19 8:53	84

プラント運転データ

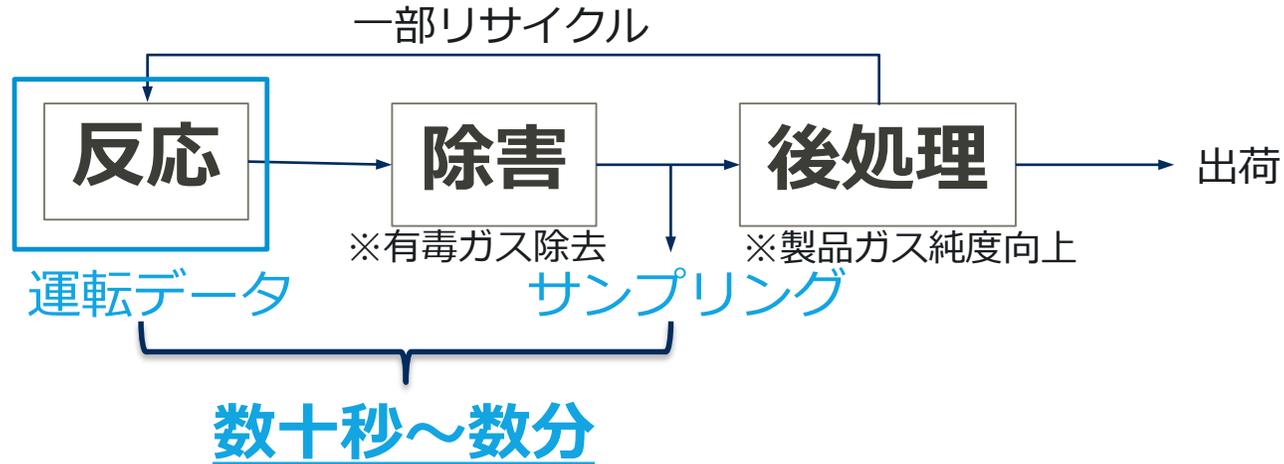
	Z1-CC0000.PV	Z1-CC0015.PV	Z1-CC0004.PV
	原料A 仕込み流量	反応器内 温度8	反応器 出口圧力
2019/5/2 8:57	18.3809	92.0560	0.6876
2019/5/2 8:58	18.3811	92.0571	0.6876
2019/5/2 8:59	18.3814	92.0582	0.6876
2019/5/2 9:00	18.3817	92.0593	0.6876
2019/5/3 10:36	18.3298	92.1694	0.6900
2019/5/3 10:37	18.3296	92.1709	0.6900
2019/5/3 10:38	18.3294	92.1724	0.6900
2019/5/3 10:39	18.3292	92.1739	0.6900
2019/5/4 10:04	18.3780	91.7240	0.6890
2019/5/4 10:05	18.3782	91.7240	0.6890
2019/5/4 10:06	18.3785	91.7240	0.6890

データセット

		Z1-CC0000.PV	Z1-CC0015.PV	Z1-CC0004.PV
	純度	原料A 仕込み流量	反応器内 温度8	反応器 出口圧力
2019/5/2 8:58	85	18.3811	92.0571	0.6876
2019/5/3 10:38	85	18.3294	92.1724	0.6900
2019/5/4 10:05	82	18.3782	91.7240	0.6890
2019/5/5 12:51	81	18.4584	92.5091	0.6911
2019/5/6 13:02	84	18.3764	92.8014	0.6901
2019/5/7 9:05	83	18.3958	92.5362	0.6889
2019/5/8 9:45	83	13.8530	90.7071	0.6902
2019/5/9 9:07	83	13.8373	93.2145	0.6899
2019/5/10 9:21	84	9.1676	90.6594	0.6894
2019/5/11 9:16	86	-0.0037	39.0421	0.0727
2019/5/18 11:08	83	13.9159	90.0084	0.6904
2019/5/19 8:53	84	20.6946	94.4640	0.6905

より純度データの予測に使えるような特徴量は何か？

◆本発表で取り扱うガス製造プロセス



◆サンプリング作業（1日15分 × 1年間 × 20件以上）



トータルで数分~十分前後のタイムラグがある事を特徴量に盛り込みたい



②純度データ記録時刻からさかのぼってプロセスデータを取得。演算して特徴量に落とし込むケース

純度データ

	純度
2019/5/2 8:58	85
2019/5/3 10:38	85
2019/5/4 10:05	82
2019/5/5 12:51	81
2019/5/6 13:02	84
2019/5/7 9:05	83
2019/5/8 9:45	83
2019/5/9 9:07	83
2019/5/10 9:21	84
2019/5/11 9:16	86
2019/5/18 11:08	83
2019/5/19 8:53	84

プラント運転データ

	Z1-CC0000.PV	Z1-CC0015.PV	Z1-CC0004.PV
	原料A	反応器内	反応器
	仕込み流量	温度8	出口圧力
2019/5/3 10:08	18.3350	92.1511	0.6900
2019/5/3 10:09	18.3348	92.1513	0.6900
2019/5/3 10:10	18.3346	92.1515	0.6900
2019/5/3 10:11	18.3344	92.1517	0.6900
2019/5/3 10:12	18.3342	92.1519	0.6900
2019/5/3 10:13	18.3340	92.1521	0.6900
	平均	平均	平均
	最大	最大	最大
	最小	最小	最小
	分散	分散	分散
2019/5/3 10:33	18.3300	92.1500	0.6900
2019/5/3 10:34	18.3300	92.1500	0.6900
2019/5/3 10:35	18.3300	92.1500	0.6900
2019/5/3 10:36	18.3298	92.1694	0.6900
2019/5/3 10:37	18.3296	92.1709	0.6900
2019/5/3 10:38	18.3294	92.1724	0.6900

前30分間

②純度データ記録時刻からさかのぼってプロセスデータを取得。 演算して特徴量に落とし込むケース

データセット

	純度	原料A仕込み流量(前30分)				反応器内温度8(前30分)				反応器出口圧力(前30分)			
		平均値	最大値	最小値	分散	平均値	最大値	最小値	分散	平均値	最大値	最小値	分散
2019/5/2 8:58	85	18.5578	18.7883	18.4448	0.0019	92.7285	95.3010	88.8517	0.0093	0.6958	0.7052	0.6958	0.0000
2019/5/3 10:38	85	18.3998	19.1548	17.7226	0.0024	92.8410	96.9716	92.5115	0.0098	0.6949	0.7079	0.6642	0.0001
2019/5/4 10:05	82	18.4379	18.7940	18.2590	0.0021	92.3246	95.3303	91.8141	0.0111	0.6922	0.7367	0.6851	0.0000
2019/5/5 12:51	81	18.4966	19.1881	18.3586	0.0026	93.4637	95.0270	93.1197	0.0103	0.6924	0.7178	0.6903	0.0001
2019/5/6 13:02	84	18.4592	19.3746	18.0827	0.0021	92.9569	95.1795	89.0525	0.0104	0.6911	0.7361	0.6863	0.0001
2019/5/7 9:05	83	18.5027	19.1136	18.3366	0.0026	92.5624	95.1008	89.2292	0.0108	0.6895	0.7359	0.6892	0.0000
2019/5/8 9:45	83	14.0302	14.4318	13.6568	0.0023	90.7429	95.4485	86.2034	0.0113	0.6986	0.7295	0.6923	0.0000
2019/5/9 9:07	83	13.8429	14.4179	13.2729	0.0027	93.5503	93.8565	90.2223	0.0093	0.6934	0.7409	0.6920	0.0001
2019/5/10 9:21	84	9.3427	9.7874	8.9669	0.0022	91.6577	92.0428	90.1523	0.0107	0.6927	0.7093	0.6505	0.0001
2019/5/11 9:16	86	0.0564	0.7450	-0.3592	0.0019	39.8879	40.7105	37.0194	0.0090	0.0822	0.0920	0.0702	0.0001
2019/5/18 11:08	83	14.0435	14.9600	13.5882	0.0023	90.1534	92.0333	88.8322	0.0093	0.6992	0.7095	0.6664	0.0001
2019/5/19 8:53	84	20.8205	21.6651	20.3649	0.0025	94.9087	97.8284	91.9869	0.0116	0.6978	0.7122	0.6839	0.0001

前30分間の情報を取り込むことで、より良い予測ができることを狙う

- 一般的には、化学の反応理論/設備の物理的な配置などを踏まえて、現実を現すような特徴量を織り込むのが良い。
- データ解析技術を併用すれば更に良いモデルが作成できる可能性がある。
(例：主成分分析、SOM、クラスターごとにモデル作成、リッジ回帰やランダムフォレストを用いた特徴量削減など)

特徴量抽出し、データセット作成完了すれば、
次は予測モデルの作成を実施する

予測モデル作成



新規セッション

ワークスペースから
ワークスペースからデータを選択し、検証方法を指定して新規セッションを開始します

ファイルから
ファイルからデータをインポートし、検証方法を指定して新規セッションを開始します

開くファイルを選択

開く(O) キャンセル

新規セッション

データセット

ワークスペース変数

1分だけのデータセット_train 350x98 table

応答

purity double 67 .. 95

予測子

	名前	タイプ	範囲
<input type="checkbox"/>	time	datetime	不適合 >
<input checked="" type="checkbox"/>	B4FC4101 APV	double	-0.018 .. 49.1445
<input checked="" type="checkbox"/>	B4FC4101 BPV	double	-0.00207813 .. 49.1845
<input checked="" type="checkbox"/>	B4FC4201 APV	double	-0.00896167 .. 49.3126
<input checked="" type="checkbox"/>	B4FC4201 BPV	double	0.0808467 .. 49.3735
<input checked="" type="checkbox"/>	B4PC4101 APV	double	-0.0316551 .. 0.152731
<input checked="" type="checkbox"/>	B4PC4201 APV	double	-0.0284945 .. 0.151479
<input checked="" type="checkbox"/>	B4FC4101 CPV	double	-1.07625 .. 19.5258
<input checked="" type="checkbox"/>	B4FC4201 CPV	double	-0.504121 .. 17.4915
<input checked="" type="checkbox"/>	B4TI4101 APV	double	86.5855 .. 443.812
<input checked="" type="checkbox"/>	B4TI4101 BPV	double	86.9268 .. 442.38
<input checked="" type="checkbox"/>	B4TI4101 CPV	double	87.9794 .. 449.646
<input checked="" type="checkbox"/>	B4TI4101 DPV	double	87.0872 .. 426.426

すべて追加 すべて削除

データの準備方法

検証

交差検証

過適合を防止するためデータセットを分割し、それぞれの部分に対して精度の推定を行います。

交差検証の分割数: 5 分割

ホールドアウト検証

大規模データセットの場合に推奨します。

ホールドアウトの割合: 25%

検証なし

過適合の防止が行われません。

検証についての説明を読む

セッションの開始 キャンセル

トレーニングデータを読み込む

予測モデル作成



回帰学習器 - 応答プロット

新規セッション ▼ 特徴選択 PCA 複雑な木 中程度の木 粗い木 すべての木 詳細設定 並列の使用 学習 応答プロット 予測と実際のプロット 残差プロット 最小 MSE のプロット エクスポート

モデルタイプ 応答プロット

データブラウザー

履歴

1 ☆ ツリー [ドラフト] 特徴 36/36 最終更新: PCA を無効にしました

現在のモデル

モデル 1: ドラフト

モデルタイプ 事前設定: 複雑な木 最小リーフサイズ: 4 代理決定分岐: オフ

最適化オプション

データセット: 1分だけのデータセット_train 観測値: 350 サイズ: 114 kB 予測子: 36 応答: purity 検証: 5 分割交差検証

回帰学習器 - 応答プロット

新規セッション ▼ 特徴選択 PCA 開始 並列の使用 学習 応答プロット 予測と実際のプロット 残差プロット 最小 MSE のプロット エクスポート

モデルタイプ すべて クイック学習

履歴

1 ☆ ツリー 最終更新: PCA を無効にしました

すべてのクイック学習

すべての使用可能なモデルタイプに学習させます

線形回帰モデル

線形 交互作用線形 ロバスト線形 ステップワイズ線形 すべての線形

回帰木

複雑な木 中程度の木 粗い木 すべての木 最適化可能なツリー

サポートベクターマシン

線形 SVM 2次 SVM 3次 SVM 細かいガウス SVM 中程度のガウス SVM

粗いガウス SVM すべての SVM 最適化可能な SVM

ガウス過程回帰モデル

有理二次 二乗指数 Matern 5/2 指数 すべての GPR モデル

データセット: 1分だけのデータセット

プロット

● 真

スタイル

● マーカー ○ 箱ひげ図 カテゴリが多すぎます

X 軸

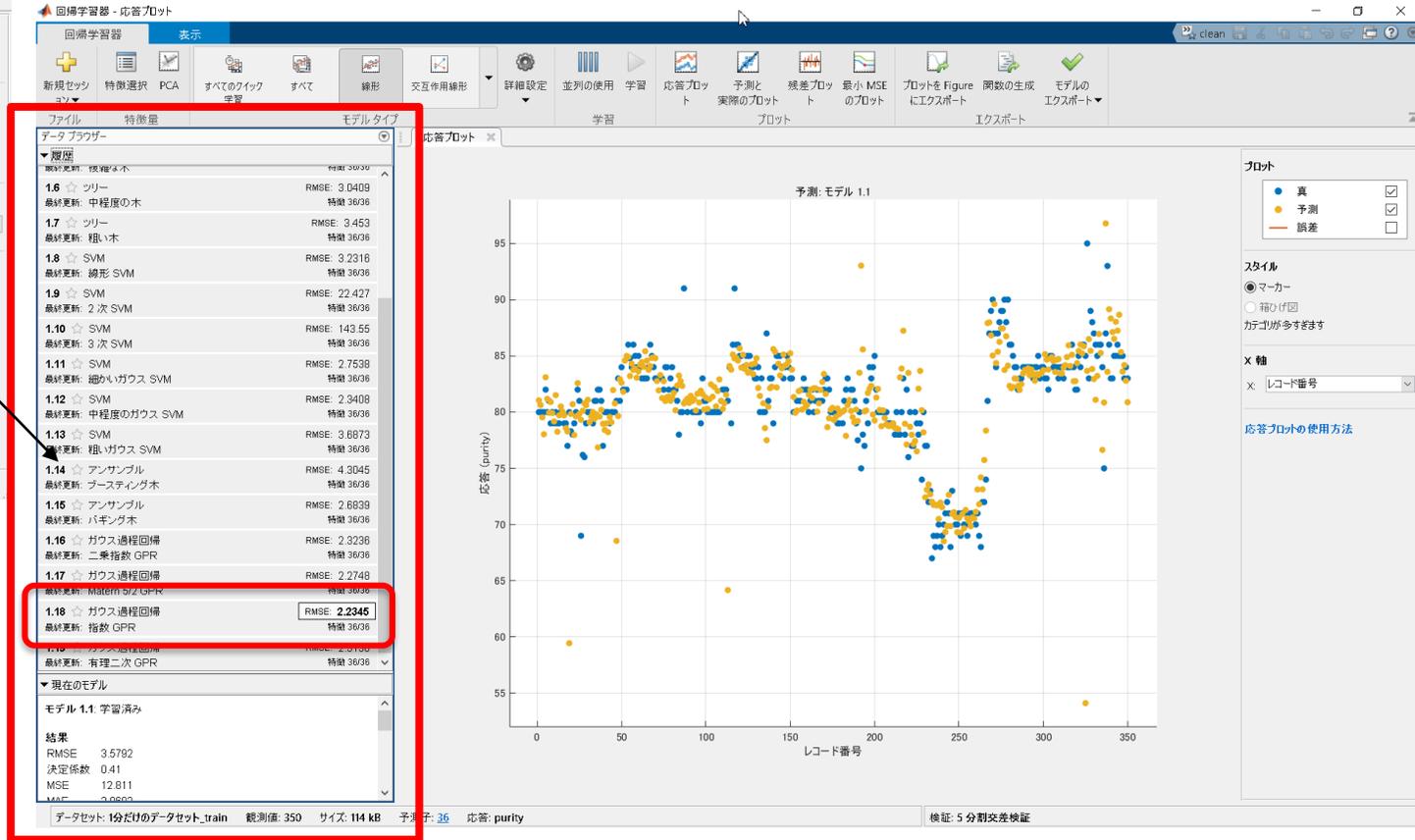
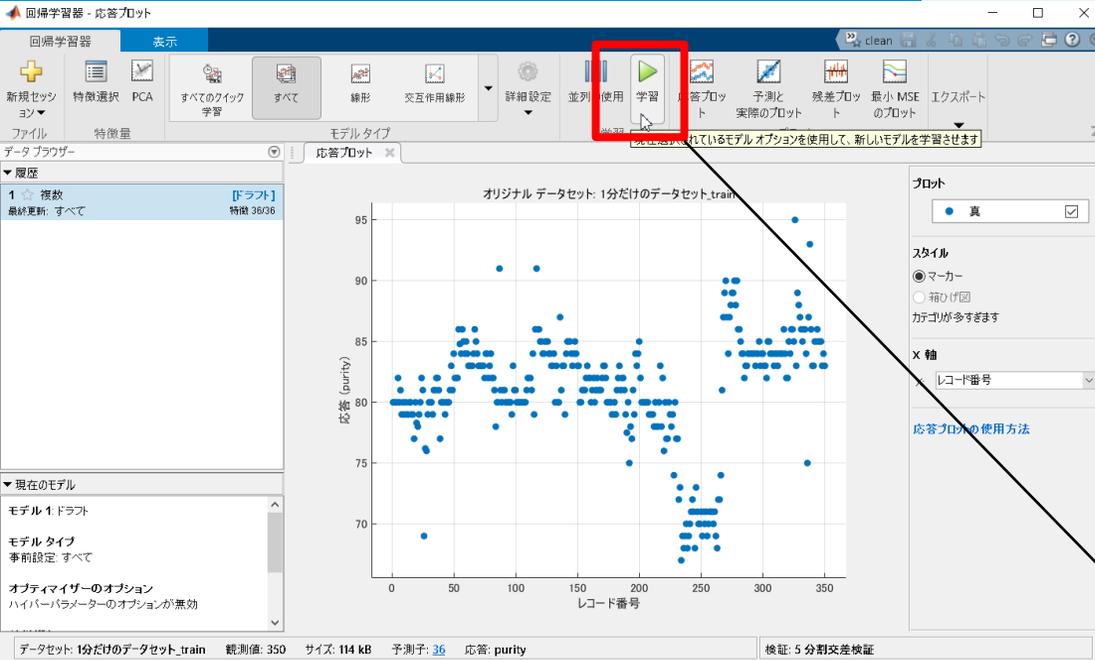
X: レコード番号

応答プロットの使用方法

データセット: 1分だけのデータセット_train

観測値: 350 サイズ: 114 kB 予測子: 36 応答: purity 検証: 5 分割交差検証

予測モデル作成



代表的な機械学習手法を複数種類試し、最も精度が高いものを示してくれる

予測モデル作成 (検証)



回帰学習器 - 応答プロット

モデルタイプ: 応答プロット

予測: モデル 1.18

RMSE: 2.2345

決定係数: 0.77

MSE: 4.9929

MAE: 1.402

データセット: 1分だけのデータセット_train 観測値: 350 サイズ: 114 kB 予測子: 36 応答: purity

検証: 5 分割交差検証

モデルのエクスポート

現在の学習済みモデルを次の名前でのワークスペースにエクスポートします: GP_Model

OK キャンセル

プロット

スタイル

● マーカー

○ 箱ひげ

カテゴリが多すぎます

X 軸

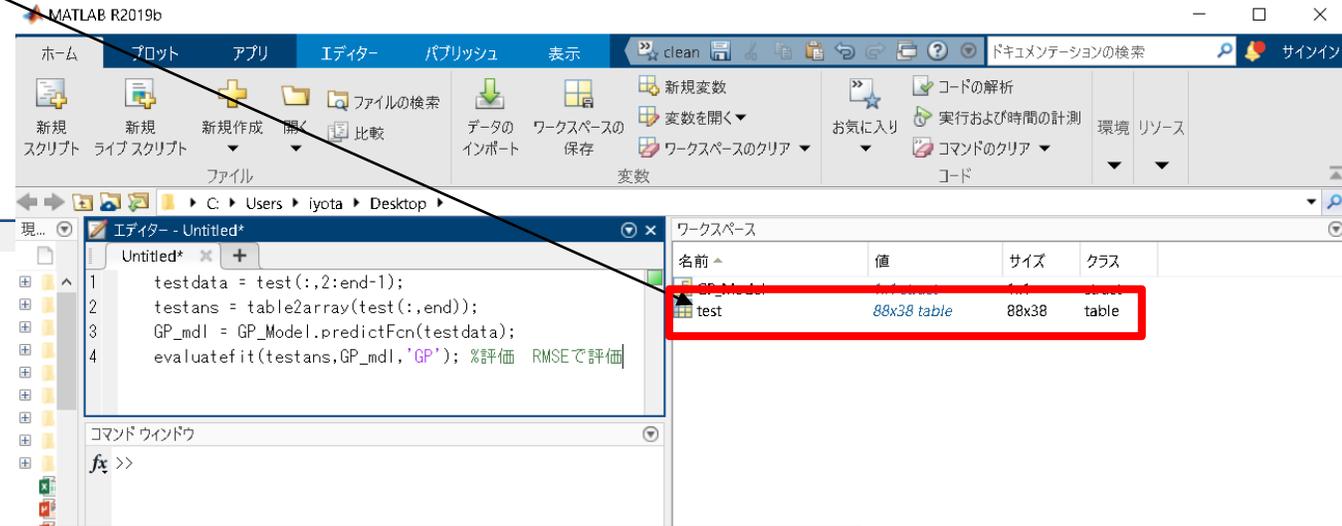
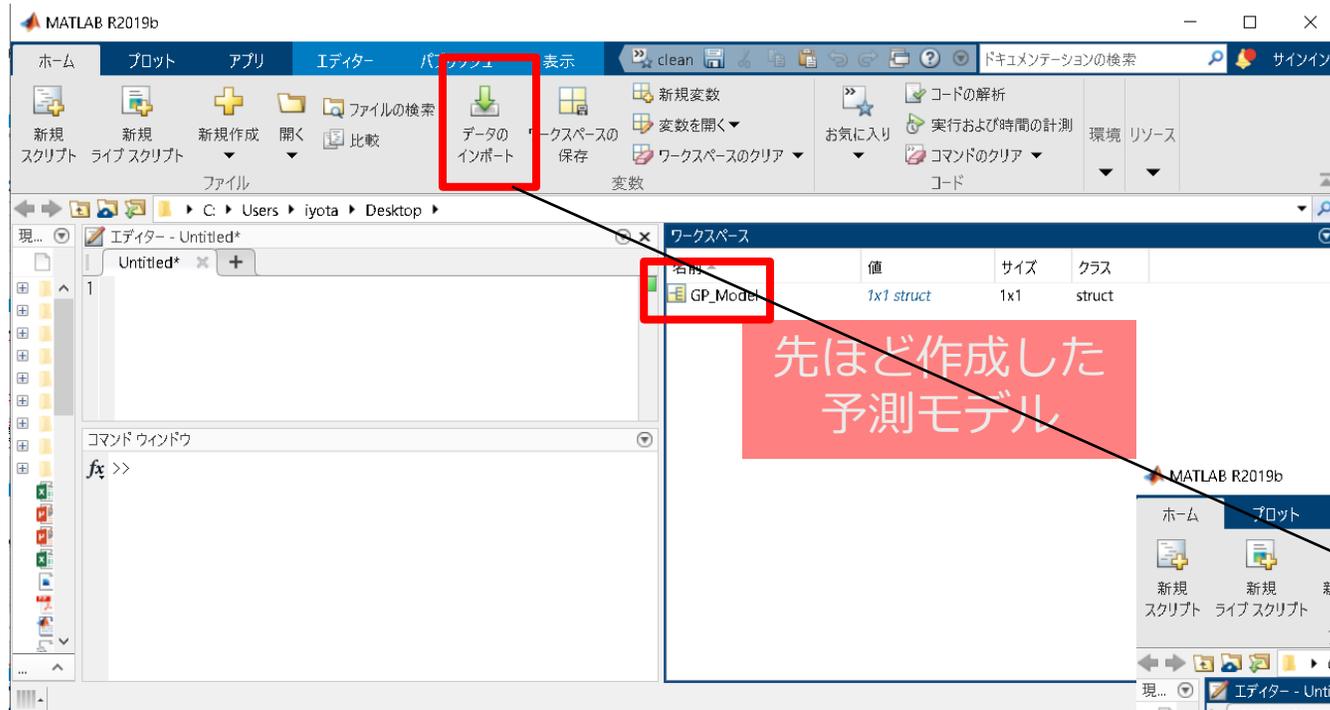
X: レコード番号

応答プロットの使用法

データセット: 1分だけのデータセット_train 観測値: 350 サイズ: 114 kB 予測子: 36 応答: purity

WinShot

検証: 5 分割交差検証



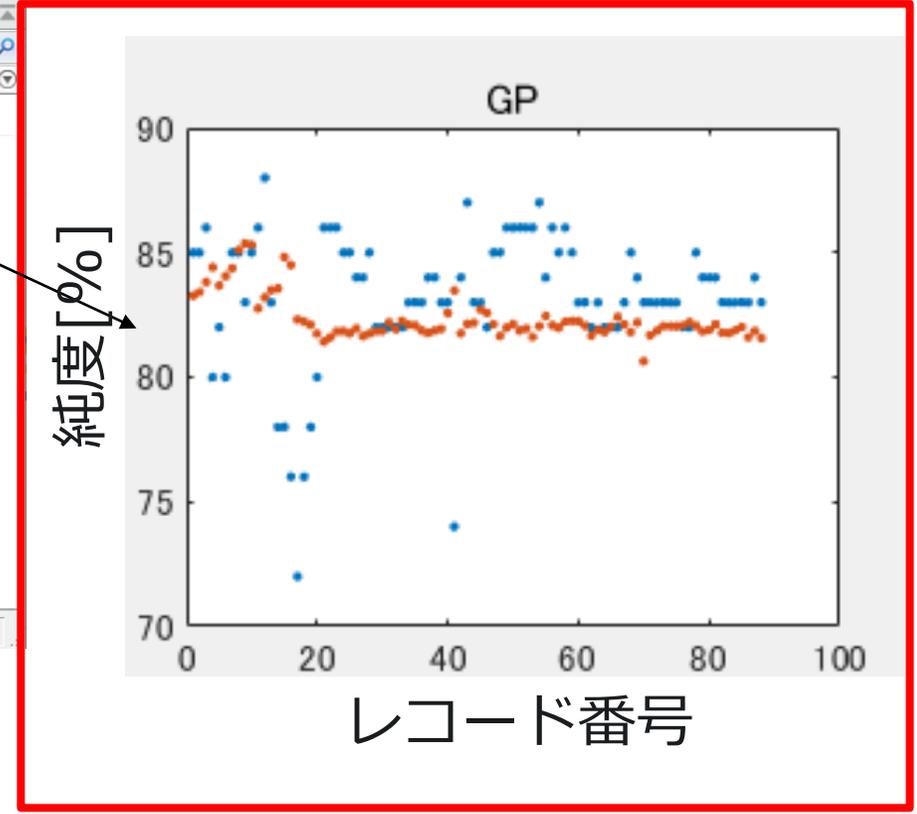
テストデータを読み込む

The screenshot shows the MATLAB R2019b environment. The '実行' (Run) button in the toolbar is highlighted with a red box. The code editor contains the following MATLAB code:

```
1 testdata = test(:,2:end-1);  
2 testans = table2array(test(:,end));  
3 GP_md1 = GP_Model.predictFcn(testdata);  
4 evaluatefit(testans,GP_md1,'GP'); %評価 RMSEで評価
```

The workspace window on the right shows the following variables:

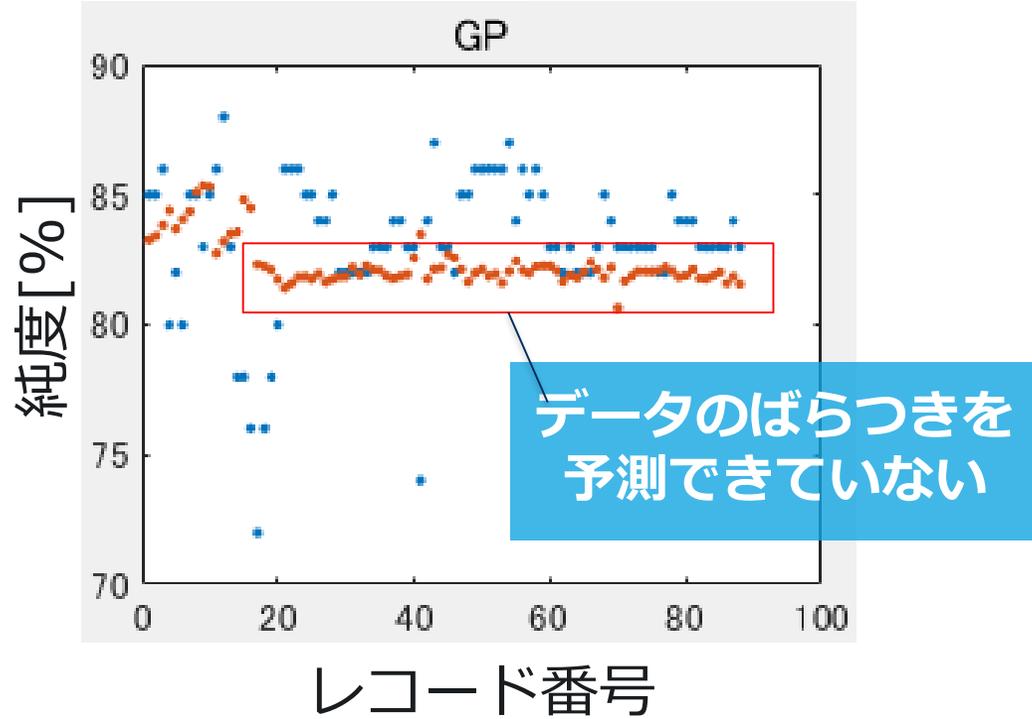
名前	値	サイズ	クラス
GP_md1	88x1 double	88x1	double
GP_Model	1x1 struct	1x1	struct
test	88x38 table	88x38	table
testans	88x1 double	88x1	double
testdata	88x36 table	88x36	table



数行のコードを書き、実行すると
テストデータの予測精度が算出される

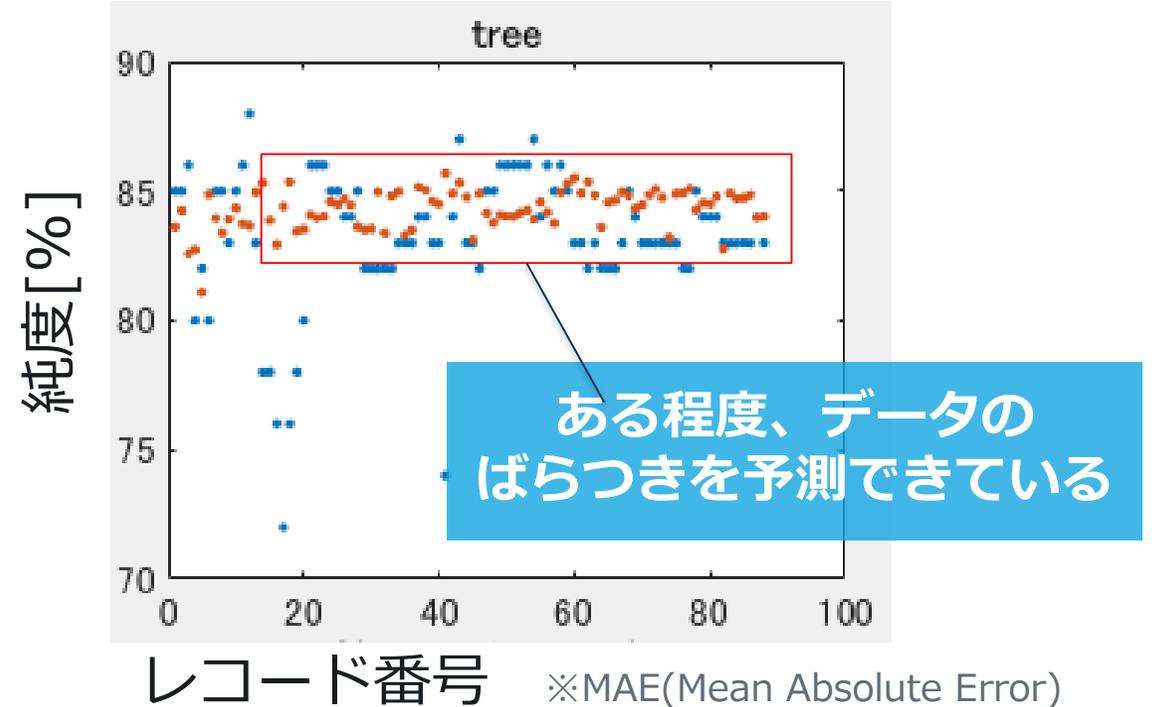
①特徴量：純度データ記録時点のみ

MAE : 2.89



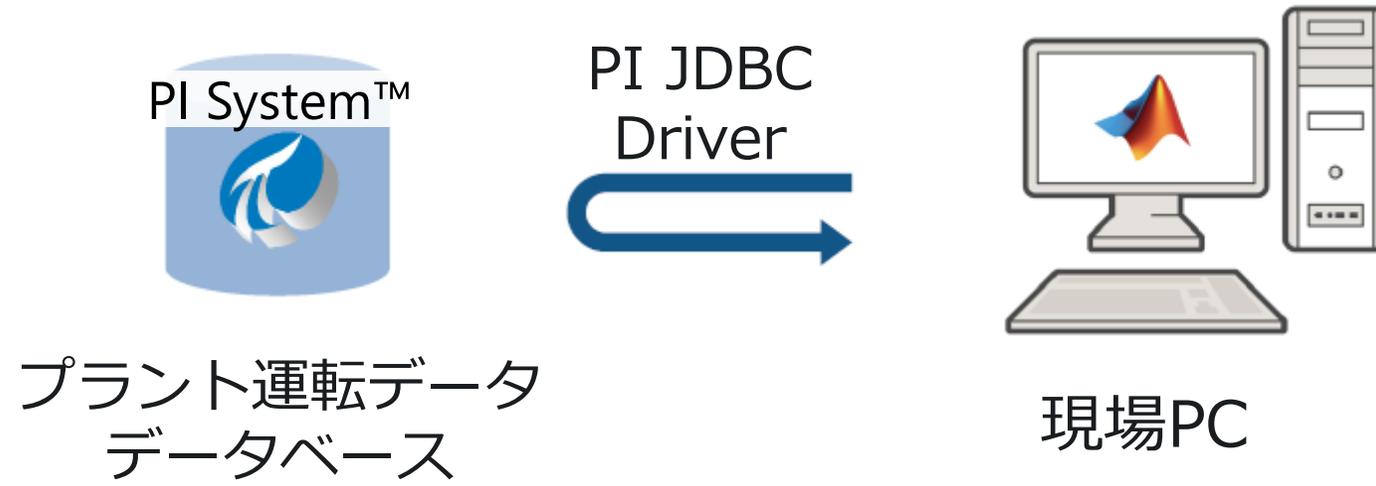
②特徴量：純度データ記録時点から30分前まで含む

MAE : 2.66



※MAE(Mean Absolute Error)
= 実測値と予測値の差の絶対値の平均値

特徴量を純度データ取得の30分前からのものにするすることで
予測精度向上が見られた



- ① 欲しいデータを取得するためのSQL文をMATLAB上に記入
- ② PI JDBC Driverを介して、PIデータベースよりデータを取得
- ③ 帰ってきたデータをMATLAB上で処理し、予測値を出力

データベースに接続

```
conn = database('PI', 'TESTSERVER¥User', 'Password')
```

SQL文でデータ取得⇒結合

```
selectquery(1)="SELECT value FROM piarchive.piavg WHERE tag = 'B4-1-1-01'.PV' AND time BETWEEN '* - 30m' AND '*";|
selectquery(2)="SELECT value FROM piarchive.pimax WHERE tag = 'B4-1-1-01'.PV' AND time BETWEEN '* - 30m' AND '*";
selectquery(3)="SELECT value FROM piarchive.pimin WHERE tag = 'B4-1-1-01'.PV' AND time BETWEEN '* - 30m' AND '*";
selectquery(4)="SELECT value FROM piarchive.pistd WHERE tag = 'B4-1-1-01'.PV' AND time BETWEEN '* - 30m' AND '*";
selectquery(5)="SELECT value FROM piarchive.pisig WHERE tag = 'B4-1-1-01'.PV' AND time BETWEEN '* - 30m' AND '*";
data_fix = [];
for num = 1:136
    [data,metadata] = select(conn, selectquery(num));
    valu(num) = table2array(data);
    data_fix = horzcat(data_fix,valu(num));
end
T = data_fix; ※T=リアルタイムに取得した特徴量データ1行分
```

平均
最大
最小
分散

作成した予測モデルを適用

```
yfit = GPModel.predictFcn(T);
```

プロット

```
plot(datetime,yfit,'b-o');
ylim([65 90]);
```

比較的短いコードでソフトセンサーが実装可能

コンパイラ

スタンドアロン アプリケーション

SoftSensor.m

ソースコード (プログラム)

アプリケーション情報

dataget 1.0

追加インストーラー オプション

アプリケーションの実行に必要なファイル

GP_model.mat

予測モデル等
必要なファイル

名前

readme.txt

SoftSensor.exe

splash.png

更新日時	種類
2020/08/31 11:34	テキストドキュメント
2020/08/31 11:34	アプリケーション
2015/06/25 22:08	PNG ファイル

現場PC(MATLABライセンス無し)でも稼働するアプリが作成可能

ヒアリング



- ・ 解決したい **課題を聞く**
- ・ データの有無/
形/場所 など

データの**変動の意味**や
背景が理解できるまで
何度も実施する。
**正確な理解が良い特徴量作成
に繋がる**

データ収集



- ・ **データ収集**
- ・ 欠損データ確認

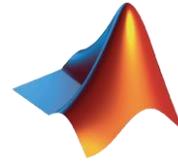
特徴量抽出



- ・ データ対応付け
- ・ **データ加工**
- ・ 予測モデル作成
できる形に
データ成形

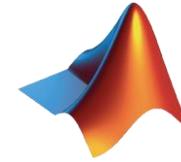
トライアンドエラー不可避
目的変数をよく表す特徴量
を作成できれば、**高性能な
ソフトセンサー**が構築できる。
どこまで突き詰めるかは
目標値、効果を見て決めるべき。

予測モデル作成



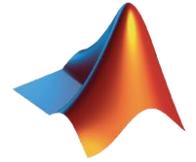
- ・ 最も**予測精度が
良い手法を
探索**する

予測モデルの リアルタイム化



- ・ データベースと接続
- ・ **リアルタイムに
予測値をプロット**する

実装(アプリ化)

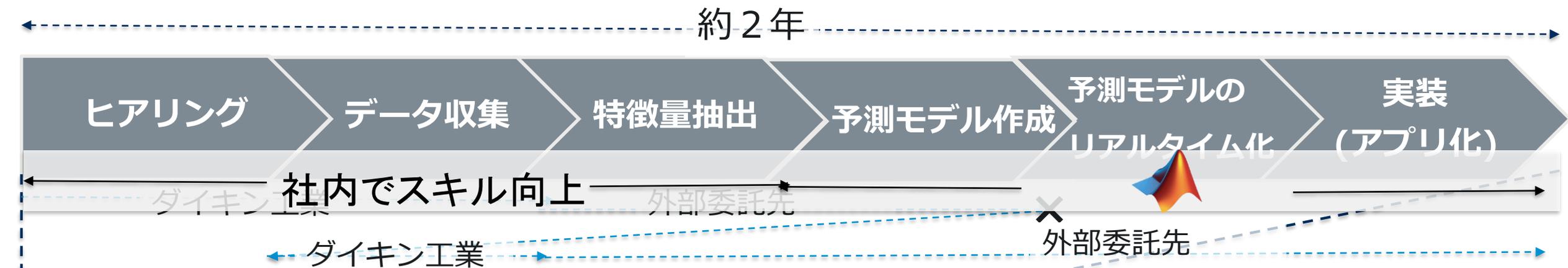


- ・ MATLAB**ライセンス
が無い環境でも動く
アプリ**を作成

MATLAB使用により大幅な
効率化が見込める。
Mathworks社の熟練エンジニア
のサポートを受けることで、
新人でもハイレベルで業務が回せる

1. 会社紹介
2. 背景情報/技術紹介
3. ソフトセンサー実装事例
4. 学んだこと/所感

過去のプロジェクト

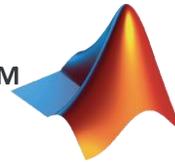


- 特徴量抽出を外部委託先に丸投げしていたためデータの背景やデータ変動の意味を何度も連絡し合う必要があった
- 予測モデルの選定も外部委託先にお任せしており、モデルの精度検証の正しさや予測モデルの意味理解に時間がかかってしまった
- 他者に仕事を完全に委託している、という意識になってしまい主体性に欠けた取り組みになっていた

約2か月

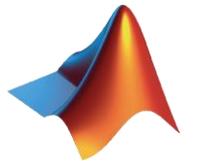


Statistics and Machine Learning Toolbox™



Database Toolbox™

MATLAB Compiler SDK™



- 予測モデル作成のトライアンドエラー期間が短かった
- 実際にリアルタイムに動くソフトセンサーを素早く実装し、製造現場の意思決定者に見せることができた
- 「手の内化」することで、担当者に主体性が生まれた

◆メリット

- Mathworks社熟練エンジニアによるテクニカルサポートがあり、技術者が「自走」できる
- 便利な各種アプリがあって、各開発工数が短縮できる
(ex.回帰学習器、分類学習器、JDBCドライバとの接続、アプリ化など)
- アプリを気軽に、オーダーメイドで作成可能
(MATLAB Production Server™ を使用して、Webブラウザから閲覧できるアプリも作成可能)
- AI人材が異動した後でもMATLABスクリプトの編集を現場で実施可能。
簡単な言語体系且つトレーニングを実施しており、初心者でも習得が容易

◆デメリット

- AI人材の技術レベルが上がれば、必要無いツールになってしまう可能性がある

◆ 「ソフトセンサーを用いた課題解決」で重要と思うポイント

- 「手の内化」をすることで担当者の主体性を上げる

- 現場の方とAI人材が密にコミュニケーションを取る

コミュニケーションが疎かだった場合・・・

⇒現場のニーズを満たしていないツールができてしまう

⇒導入時には使ってもらえても、業務として根付かず、使い続けてもらえない
ツールができてしまう

- データの意味を、ヒアリングで深堀する

数回の情報交換だけだと、必ず共有しきれていない情報がある。

データ解析の途中で何度でも情報交換が必要。

ご清聴ありがとうございました

