



CENTRAL GLASS

硝子分野における データサイエンスの活用

硝子研究所
8研 藤井剛
3研 齋藤彰太

前年度までの取り組み

“誰でもデータサイエンスができる汎用ツール”：テーブルデータからの予測

教師データ

出力結果

IDなど

説明変数

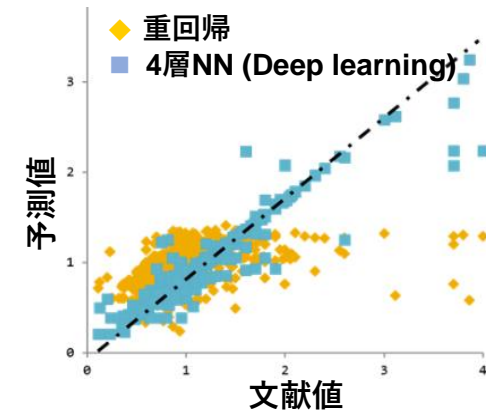
目的変数

No.	A	B	C	D	E	F	G	H	I	Q	R	S	T	U	V
	GlassNo.	DataSource	Year	GlassNo.	SiO2	B2O3	Al2O3	MgO	CaO	TiO2	ZrO2	CeO2	P2O5	H2O	FractureT
1	12630	GP02-272	Japanese	2006	67.76	0	2.33	8.68	0	1.12	0	0	0	0	0.74
2	15670	GP02-307	Japanese	2012	62	0	13	3	0	2	0	0	0	0	0.89
3	676	GI02-032	J.Ceramic	1988	71	0	1	0	0	16	0	0	0	0	0.7
4	2732	GI06-037C	US.America	1978	60	0	22.5	0	0	0	0	0	0	0	0.65
5	8683	GI06-1916	NGF>Add	2001	55.15	0	14.38	0	0	0	0	0	0	0	0.81
6	6773	GI03-156	GlassTech	1994	60	0	0	0	0	0	0	0	0	0	0.91
7	15805	GP05-307	USPatent	2012	67.2	10.6	10.6	2.9	0	0	0	0	0	0	0.9
8	9160	GI05-2071	J.Ceramic	2001	57	19	0	0	0	0	0	0	0	0	0.68
9	10541	GP03-225	USPatent	2002	66.42	0	4.76	3.44	0	1.69	0	0	0	0	0.65
10	15085	GP02-301	Japanese	2011	61.9	0	13	3	1	0.6	0	0	0	0	0.94
11	15803	GP05-307	USPatent	2012	66.2	7.6	11.3	5.3	0	0	0	0	0	0	0.9
12	2371	GI02-045	Physican	1980	85	0	0	0	0	0	0	0	0	0	0.8
13	14551	GI02-298C	J.Non-Cry	2010	66.7	0	0	0	0	0	0	0	0	0	0.74
14	4848	GI02-140C	J.Material	1991	74.35	0	0	0	0	0	0	0	0	0	0.95
15	12457	GP02-268	Japanese	2005	65.48	0	0.57	22.97	0	0.47	0	0	0	0	0.875
16	15802	GP05-307	USPatent	2012	68.13	9.79	10.39	0	0	0	0	0	0	0	0.9
17	12432	GP02-268	Japanese	2005	65.13	0	3.58	15.4	0	1.43	0	0	0	0	0.87
18	6867	GI07-1851	J.Nature/Pat	1998	76	0	0	0	0	0	0	0	0	0	0.64

機械学習

CAB

パラメータ変更は不可



予測

探索

- ・重回帰分析
- ・ランダムフォレスト
- ・スパースモデリング

- ・サポートベクターマシン
- ・勾配ブースティング木
- ・ディープラーニング

- ・決定木
- ・SGD

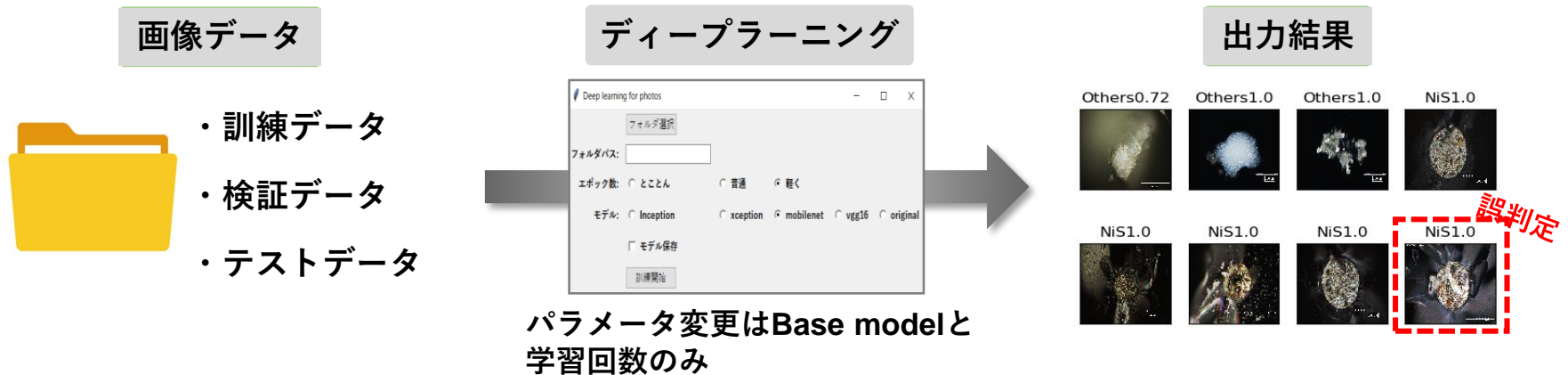
- ・グリッドサーチ
- ・ベイズ最適化

ex. ガラス組成→物性の予測

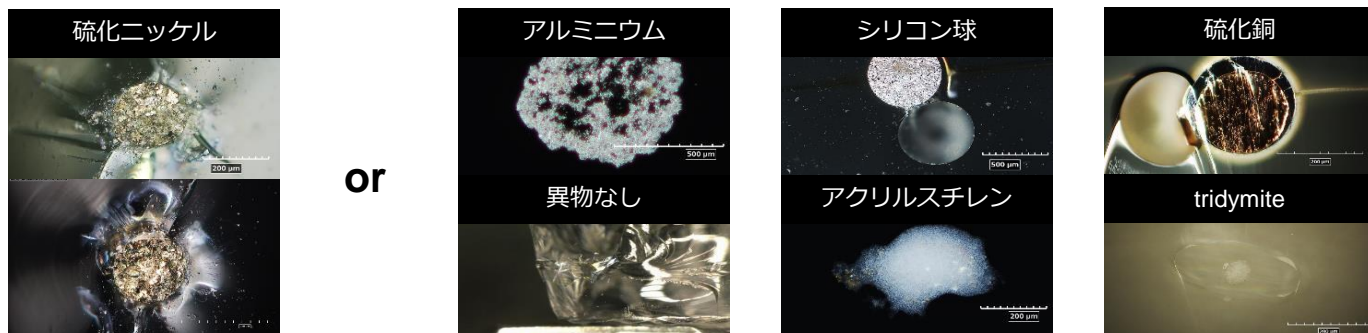
教師データに公開DB (Interglad) を用い、目的の破壊靱性値を有するガラス組成を予測

前年度までの取り組み

“誰でもデータサイエンスができる汎用ツール”：画像データからの予測



ex. 風冷強化ガラスの自爆原因分析の効率化 (2クラス分類)



今年度の取り組み



前年度までは...

汎用的なツールの作成を重要視していた

○ 機械学習に適したデータが収集できているか？
→ データ収集・選定方法、形式、数、前処理など

○ 課題に適した機械学習ができているか？
→ アルゴリズム、パラメータ、評価指標など

課題を理解している人が適切なツールを使って
初めて有益な情報が得られる

具体的な課題の抽出・解決

活用事例①

白金ポット形状最適化プロセスの効率化

活用事例②

モアレ検査における異物分類の効率化

部会の立ち上げ

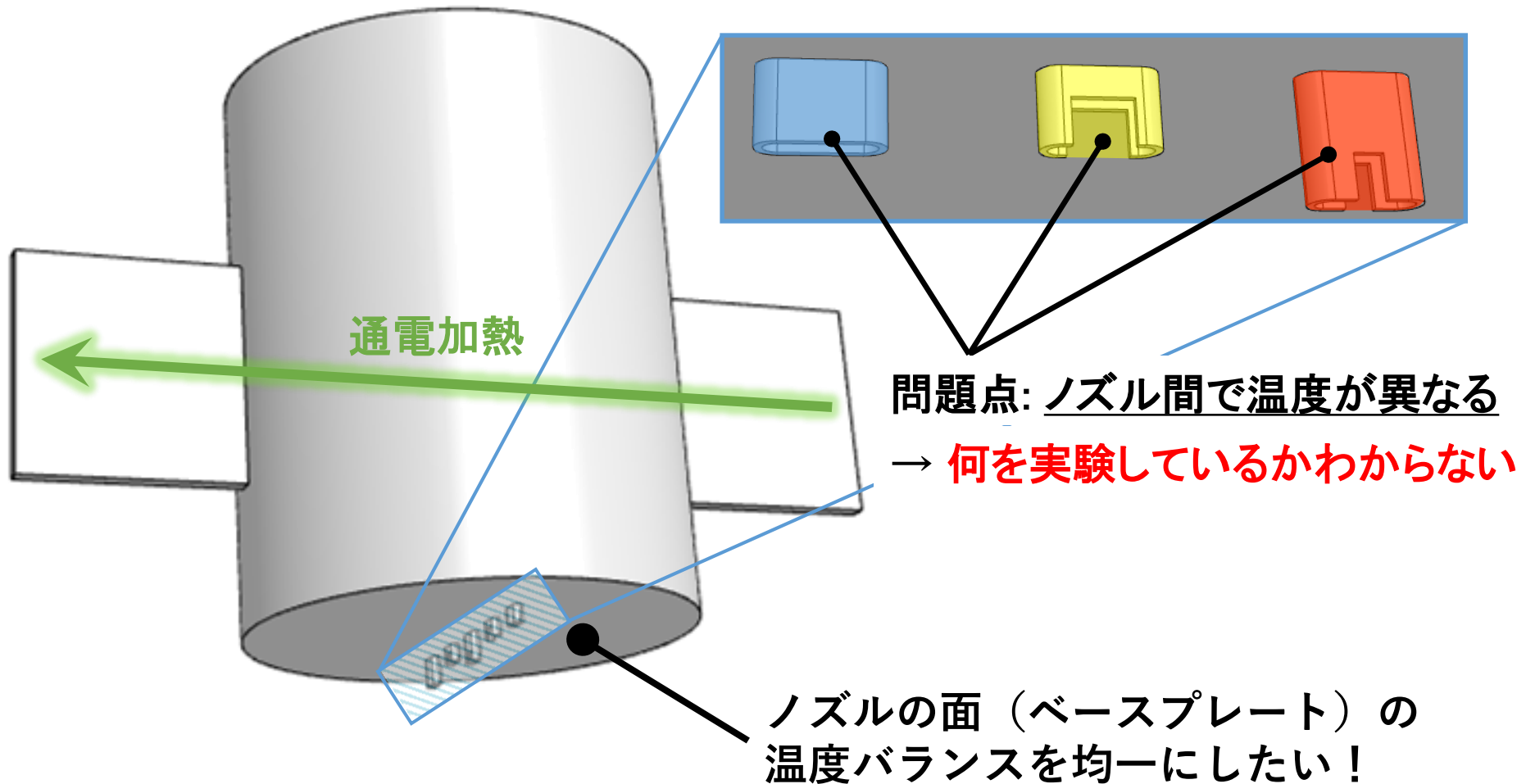
- ・ 7月末に硝子研の部会活動の一環として始動
- ・ 実務を通じたデータサイエンスの土台作りが第一の目的

活用事例①

ガラス繊維分野における
シミュレーション技術の活用
＋
機械学習による探索手法の改良

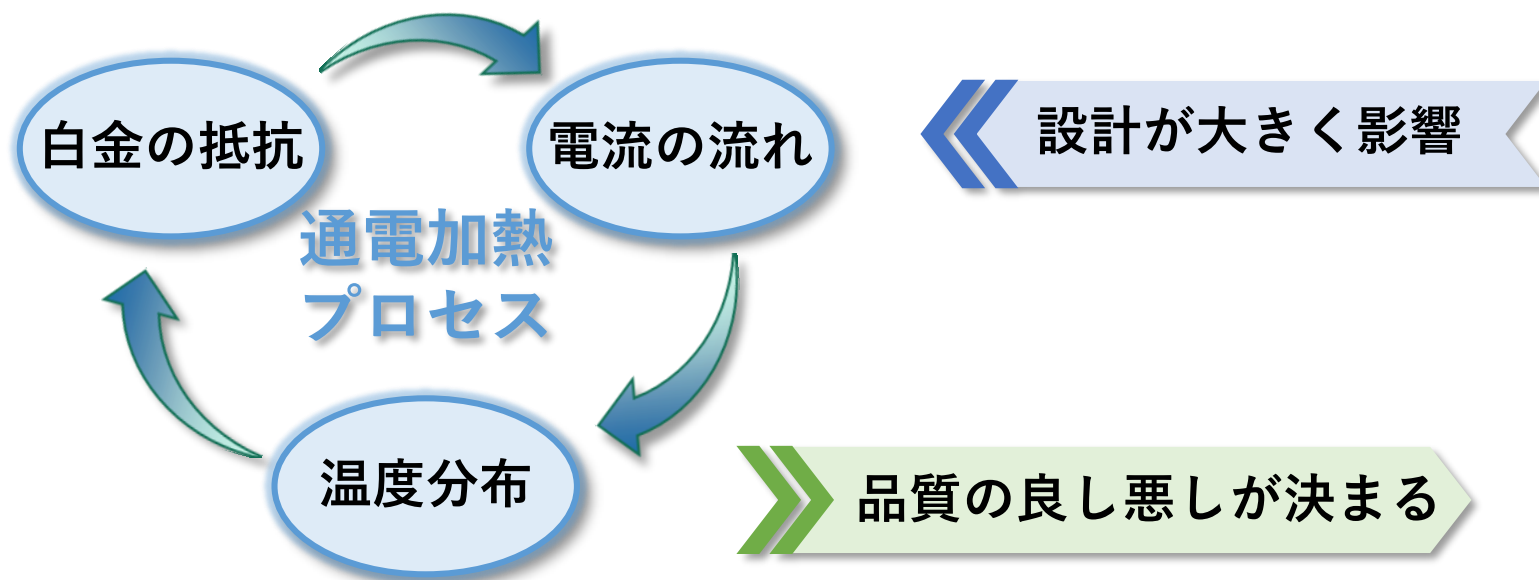
ガラス繊維の研究開発における問題

- オリジナルのノズル付き白金製ポットを製作して紡糸実験



通電加熱の難しさ

- 通電対象の形状／設計で温度バランスが大きく変わる

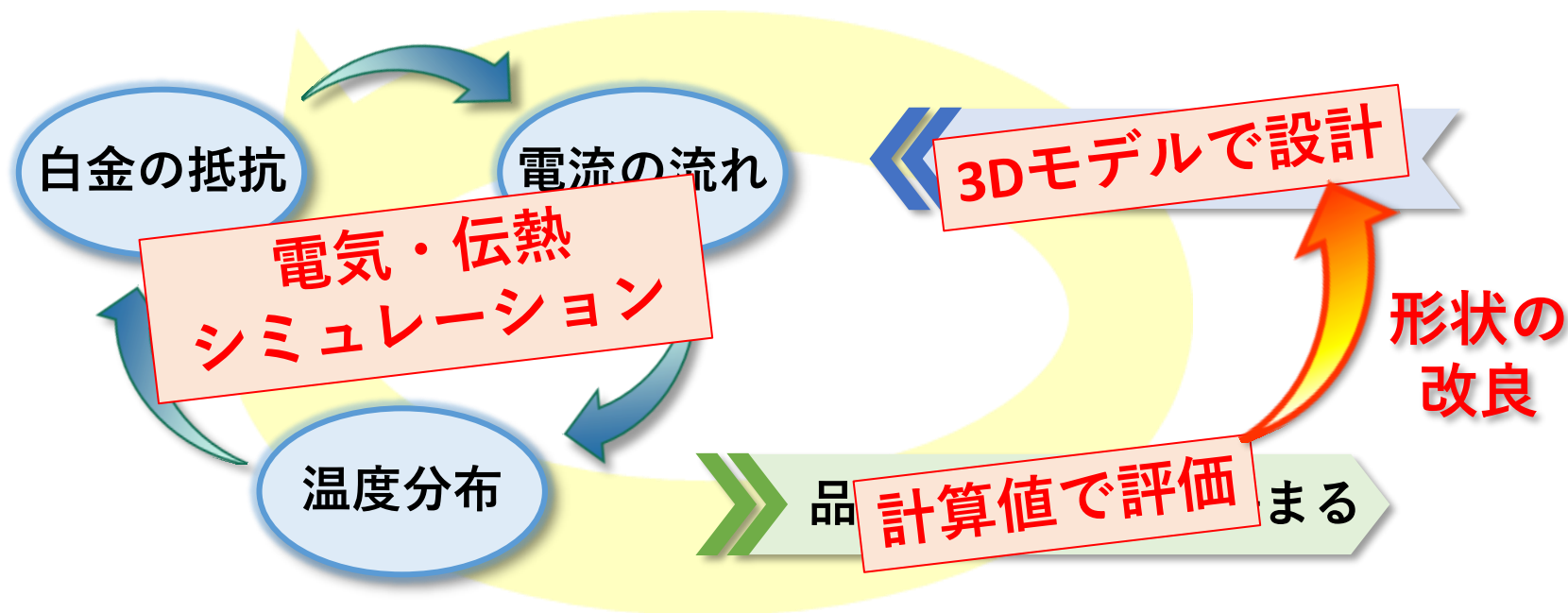


❌ 複雑なプロセス → 最適化は容易でない

❌ 製作は外注 → トライアンドエラーの最適化では莫大なコストと時間

シミュレーションの活用

- 複雑な通電加熱をシミュレーションに解かせて評価

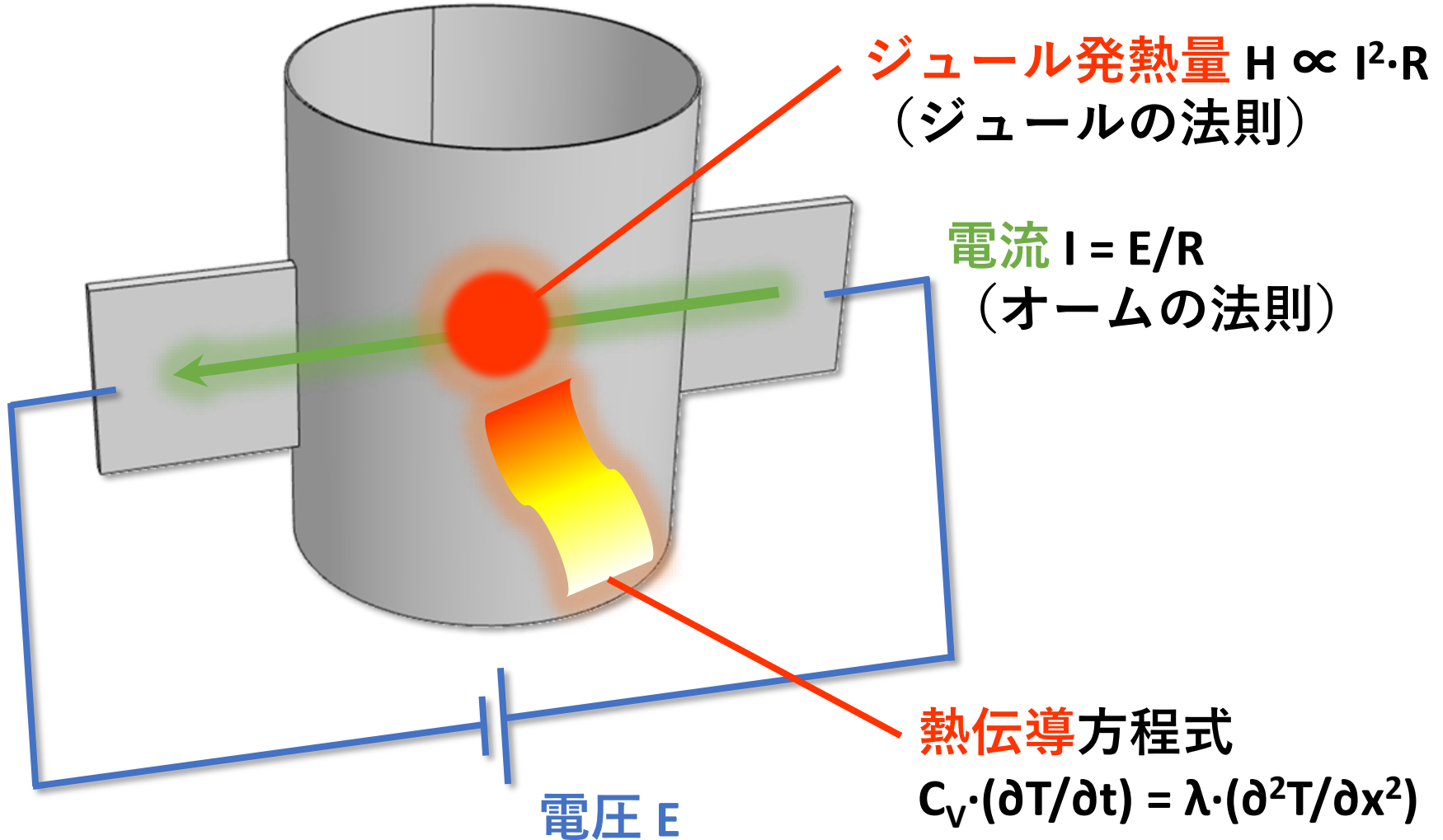


○ 形状最適化プロセスの高効率・低コスト化

使用ソフト: COMSOL Multiphysics®

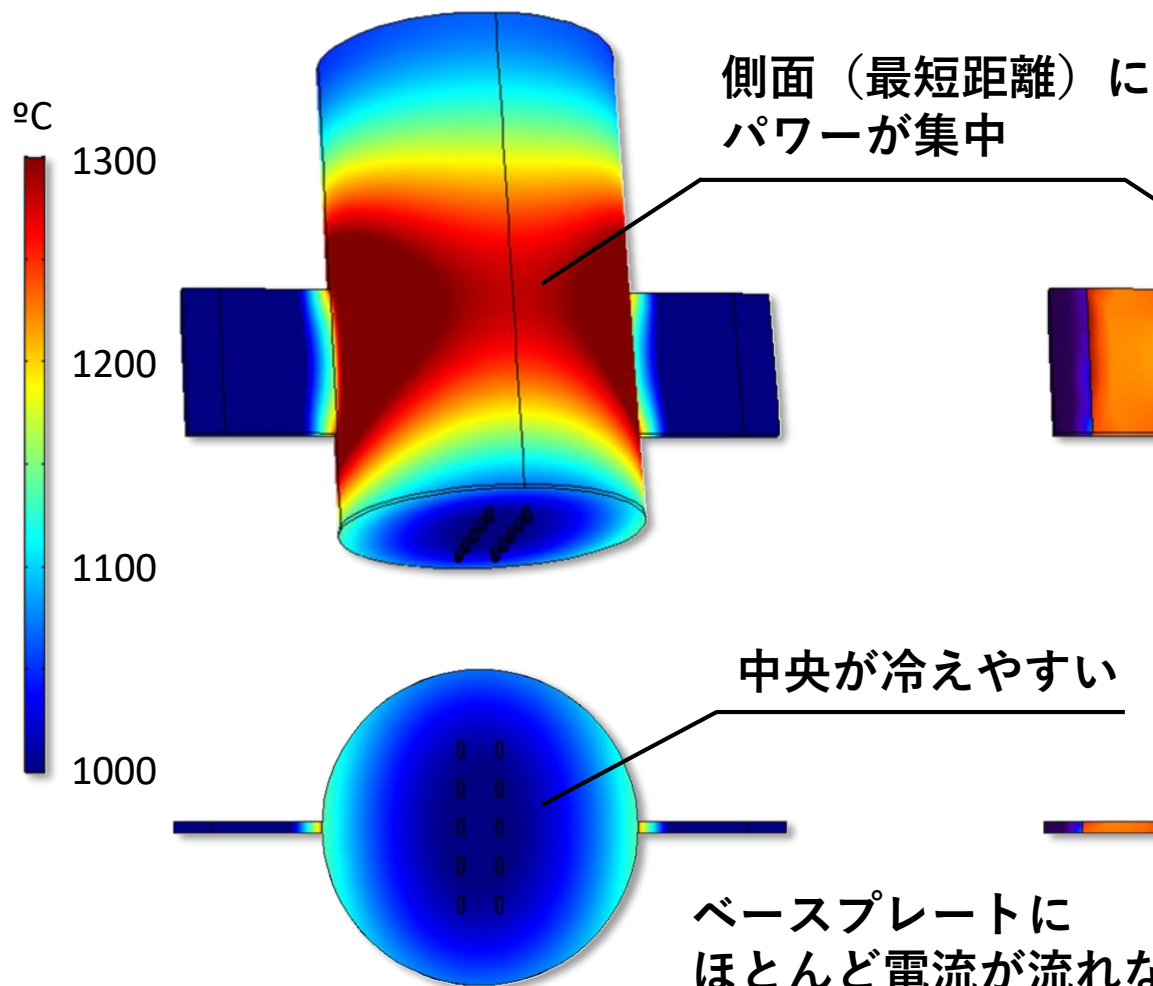
化研宇部で契約, 複数の物理現象を連成して解くことに長ける
→ 例: 電気 + 伝熱

計算モデル

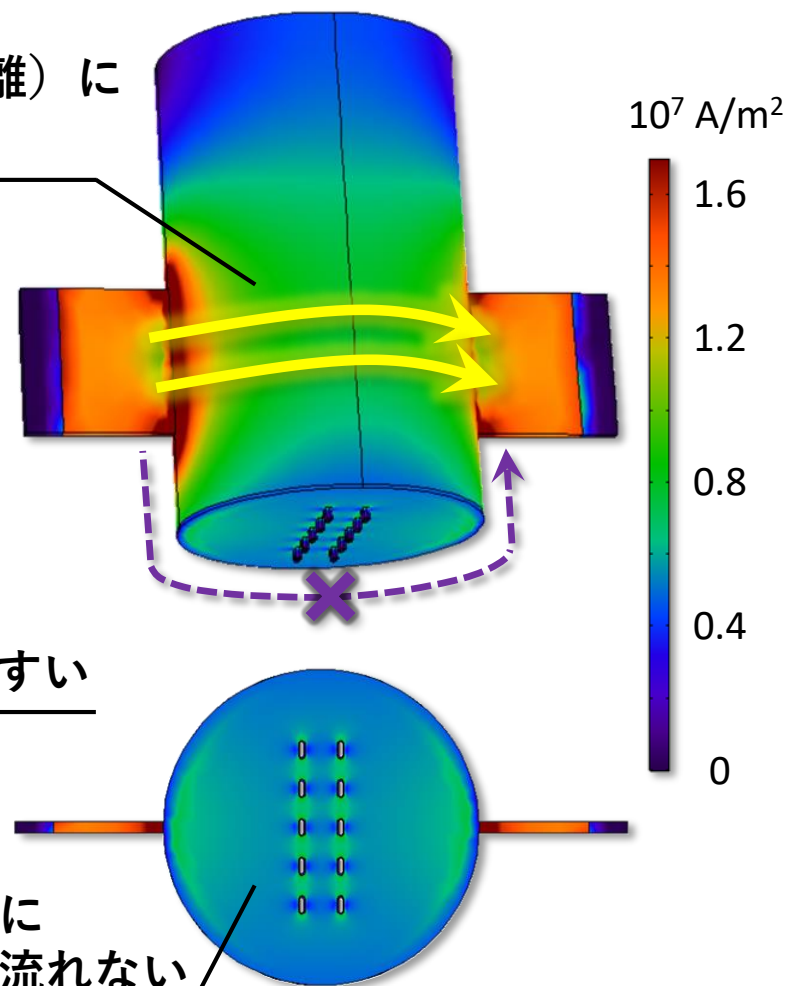


白金ポット設計の問題点

◆温度分布

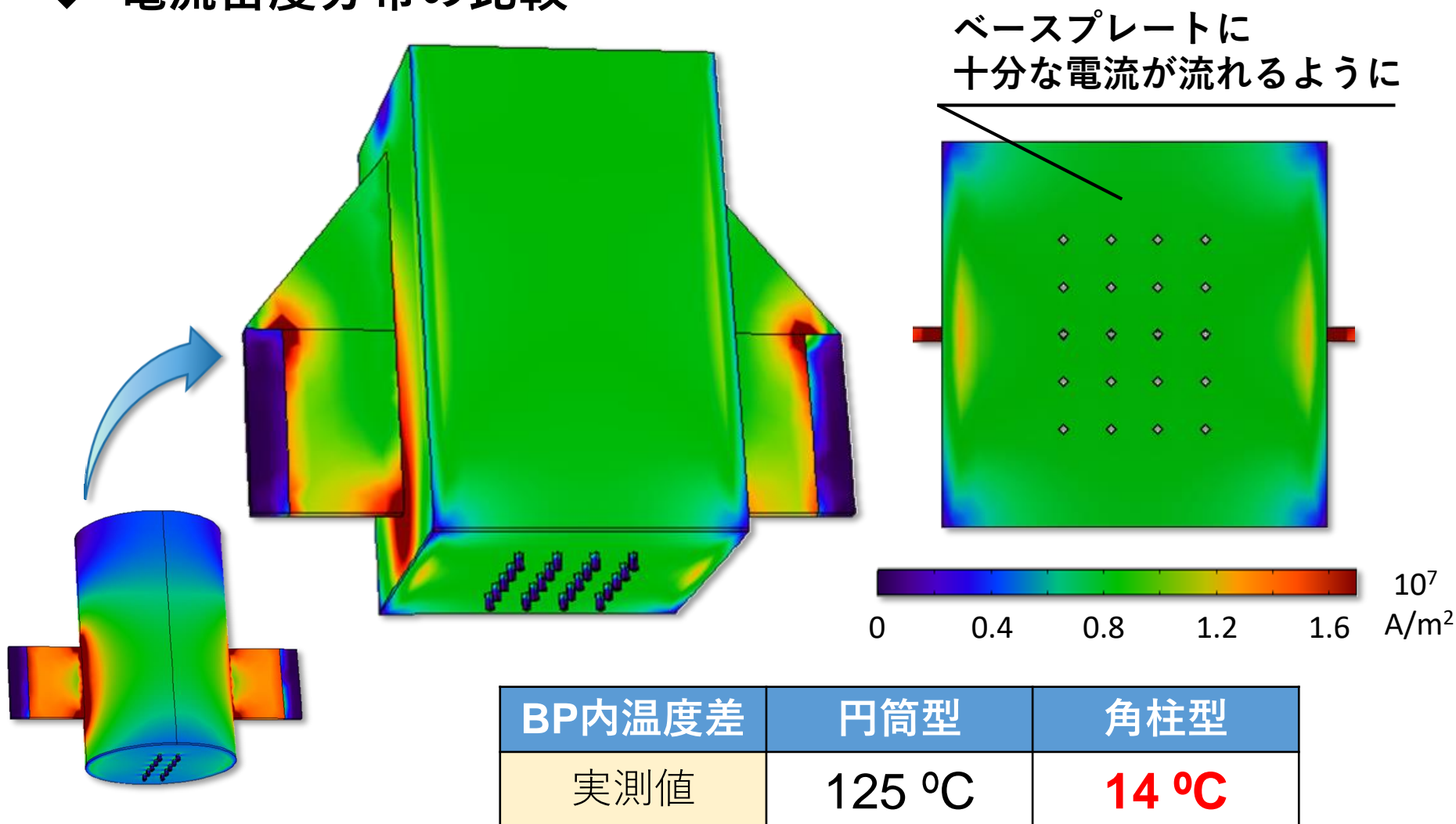


◆電流密度分布



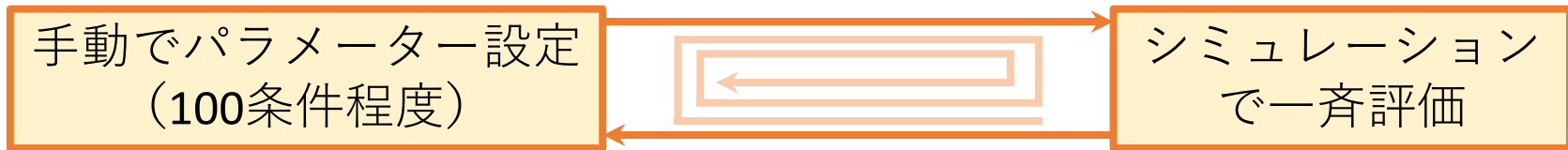
白金ポット形状の改良

◆ 電流密度分布の比較



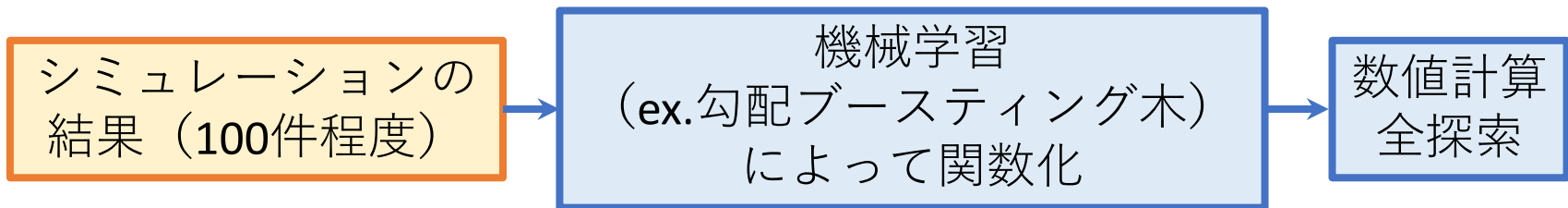
機械学習による探索の効率化

現行の最適化



- パラメーターの選び方は経験やカン → 適切？
- 計算量は100条件程度／日が限界
- 粗い全探索しかできない

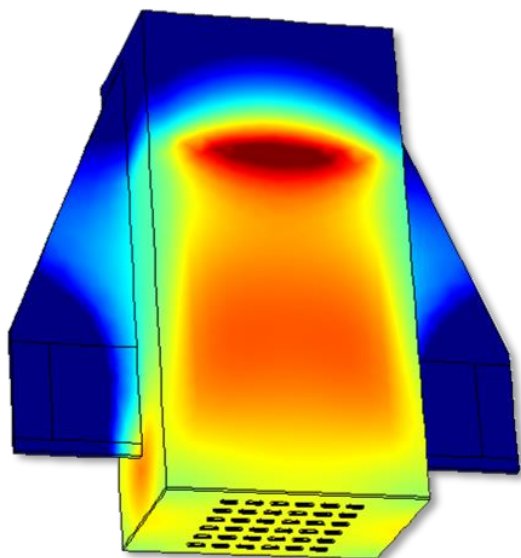
機械学習との融合



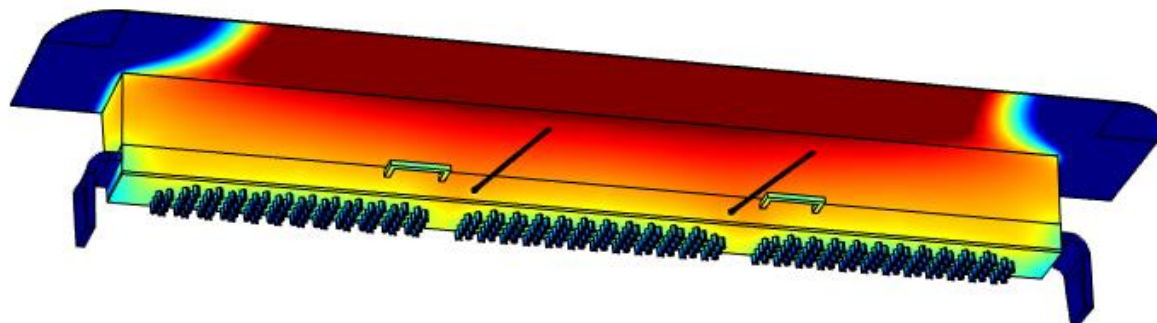
白金ポットの最適化に適用した結果

	シミュレーション のみ (144条件)	← + 機械学習 (7万条件)
目的の温度差 (計算値)	10.7 °C	6.7 °C

実験用～生産用まで幅広く活用



硝子研 実験用白金ポット



CGF 生産用大型白金ポット
(ブッシング)

従来の開発
(トライ＆エラー)

- ✗ 設計の指針なし
- ✗ どこを変更すべき？



シミュレーション＋機械学習
を活用した開発

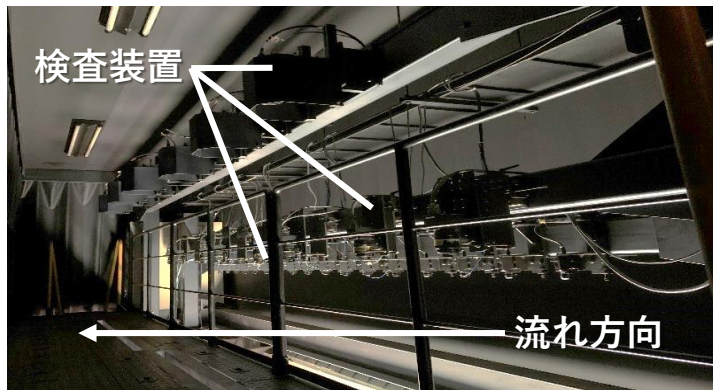
- ✓ 視覚化・手軽な変更検討
- ✓ 数回の製作で大幅改善

活用事例②

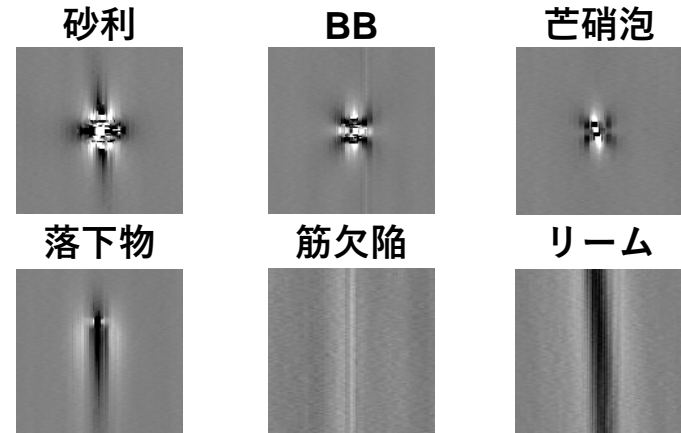
モアレ検査における異物分類の効率化

モアレ検査における従来の異物判別方法

松阪工場におけるモアレ検査
→ 素板(徐冷後)の異物検出



異物種別のモアレ画像



異物が頻発

モアレ画像を元に異物判別
実際に採断して確認



生産条件にフィードバック

(数時間後)

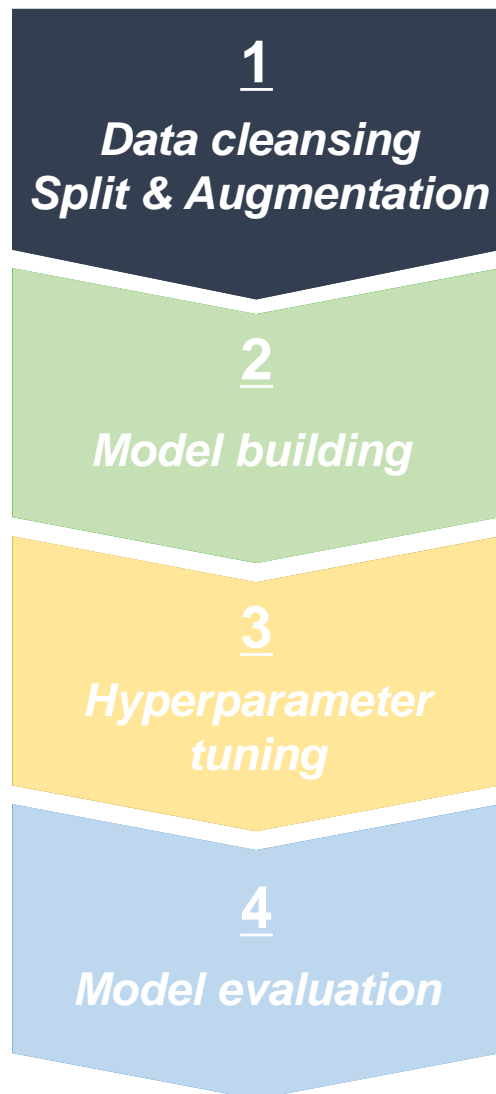


詳細な分析を分室に依頼・異物種が確定

✗ 経験値がないと目視での判別は難しい ✗ 生産条件へのフィードバックの遅さ

目的：「異物判別時間の短縮・効率化」

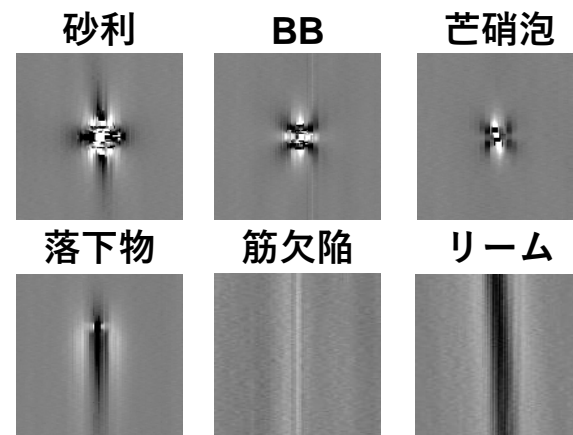
画像分類モデル構築の流れ



Data cleansing

- 品種 : MFL2
- 異物種 : 6種
- 枚数 : 約1000枚
- 画素数 : 128×128
- ファイル形式 : PNG (8bit)

*検証では4/27生産時のデータを使用



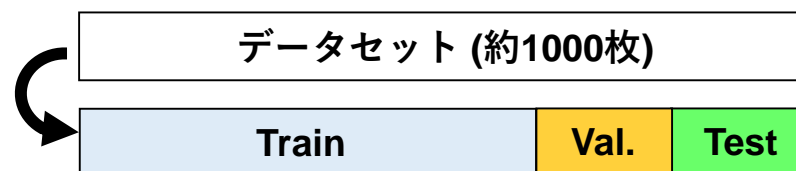
✓ 間違われやすいデータは十分確保

✗ 人間でも判定に困るデータ

✓ 画像サイズや異物の位置

✗ 誤ったラベル付け

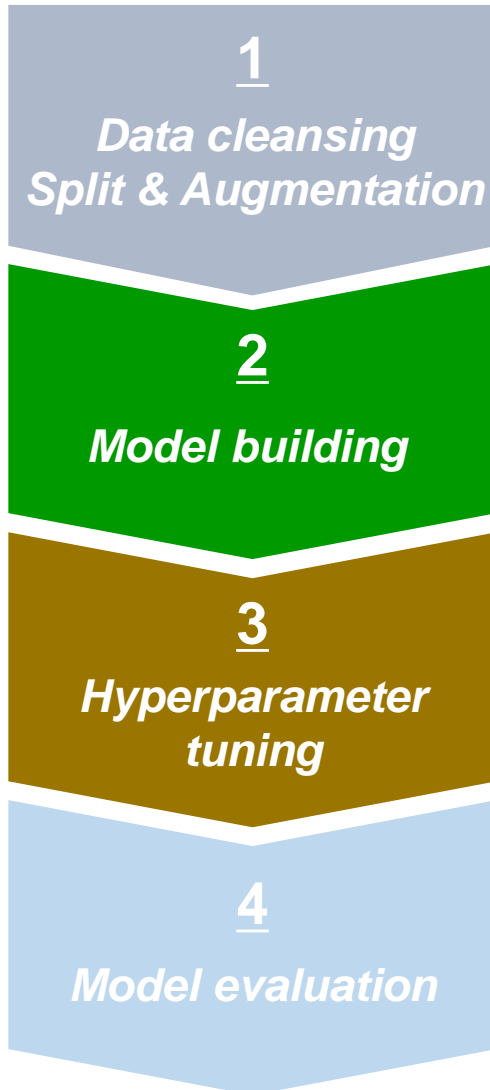
Data split & Augmentation



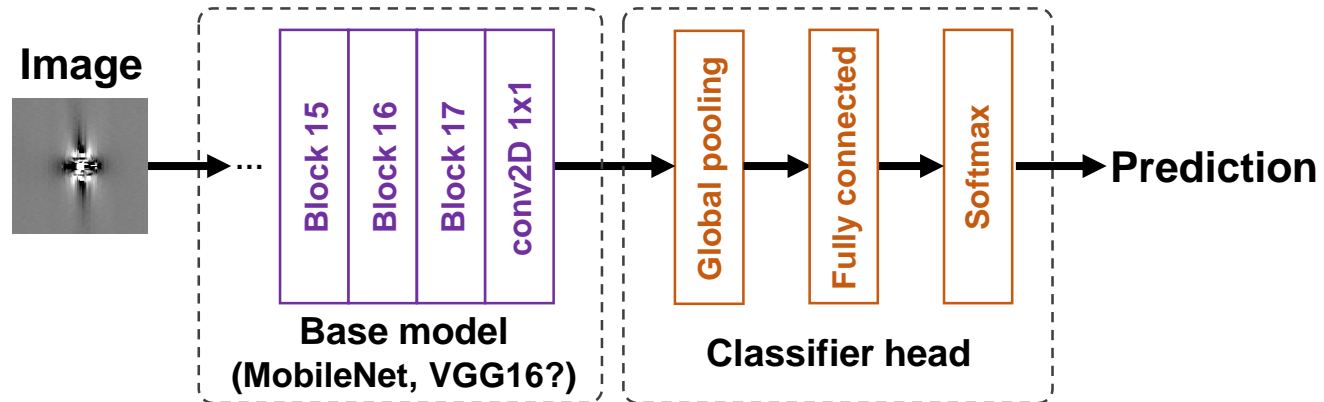
Keras

{ 画素値変換, 拡大縮小,
シアー変換, 上下左右反転,
上下左右平行移動, RGB補正

画像分類モデル構築の流れ



Model building



Hyperparameter tuning

Learning rate = 1e-3, 1e-5 ?

Batch size = 8, 32, 64 ?

Epoch = 10, 100, 1000 ?

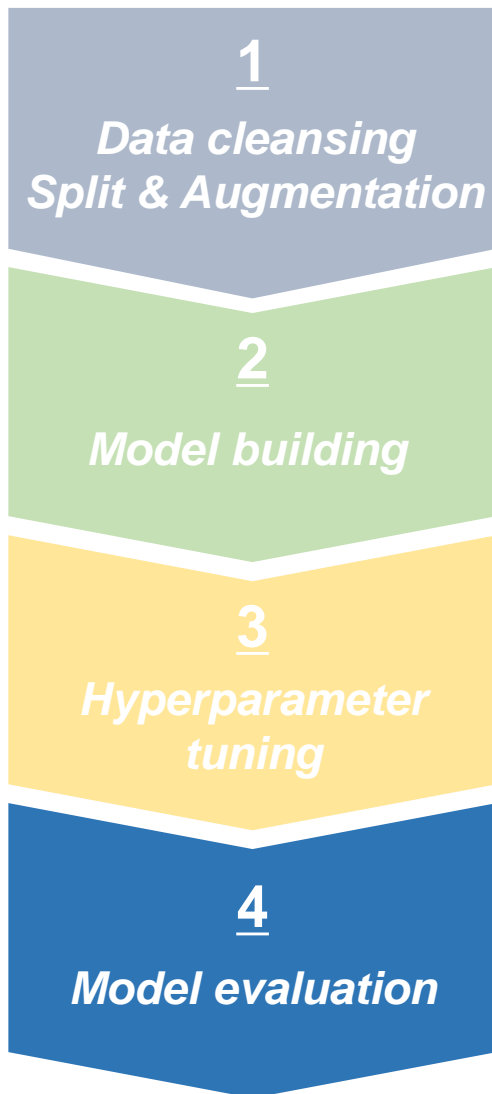
Optimizer = Adam, RMSprop, SGD ?

Activation = Softmax, Sigmoid ?

or

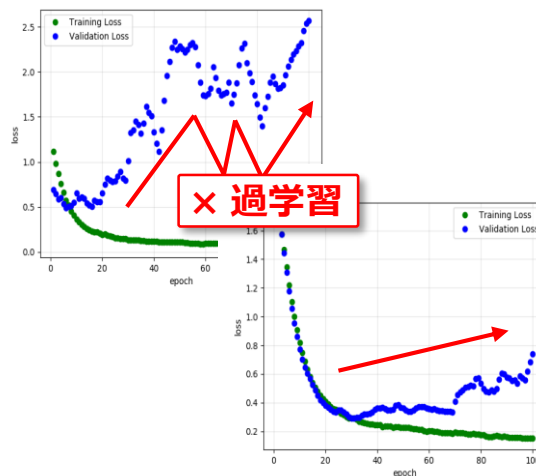
全探索、ランダムサーチ、ベイズ最適化

画像分類モデル構築の流れ



Model evaluation

Loss-curve (学習の様子)



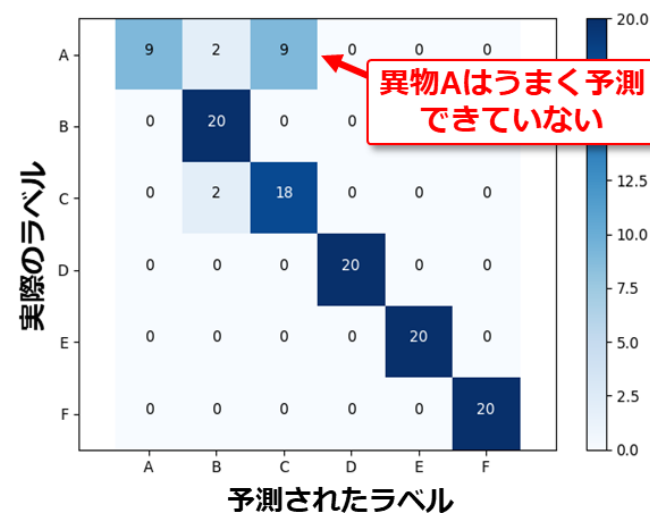
正答率 : Accuracy

全体に占める正解の割合

適合率 : Precision

分類結果がどれだけ正しいか

Confusion-matrix (混同行列)



再現率 : Recall

正解をどれだけ取り逃さなかったか

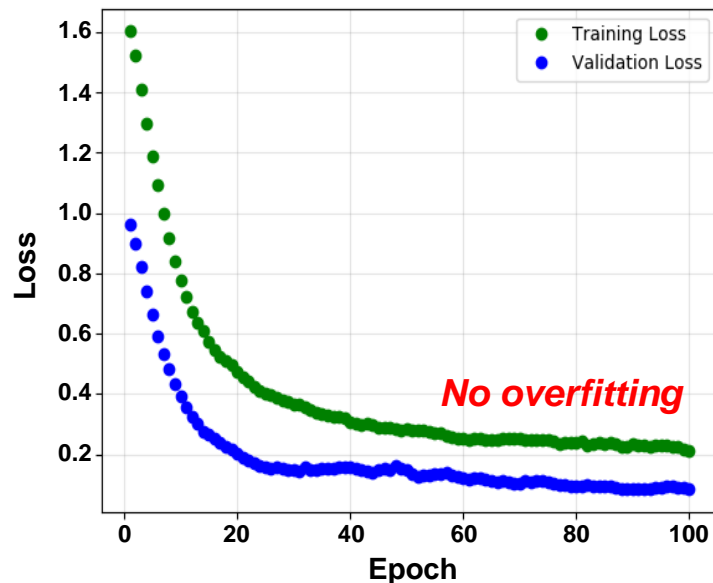
F値 : F-measure

PrecisionとRecallの良いところ取り

最適化モデルを用いた異物画像の分類結果

Model design (Best)

- Input size : 224 x 224 x 3
- Base model : MobileNet
- Activation : Softmax
- Optimizer : Adam
- Learning rate : 1e-5
- Batch size : 8
- Epoch : 100

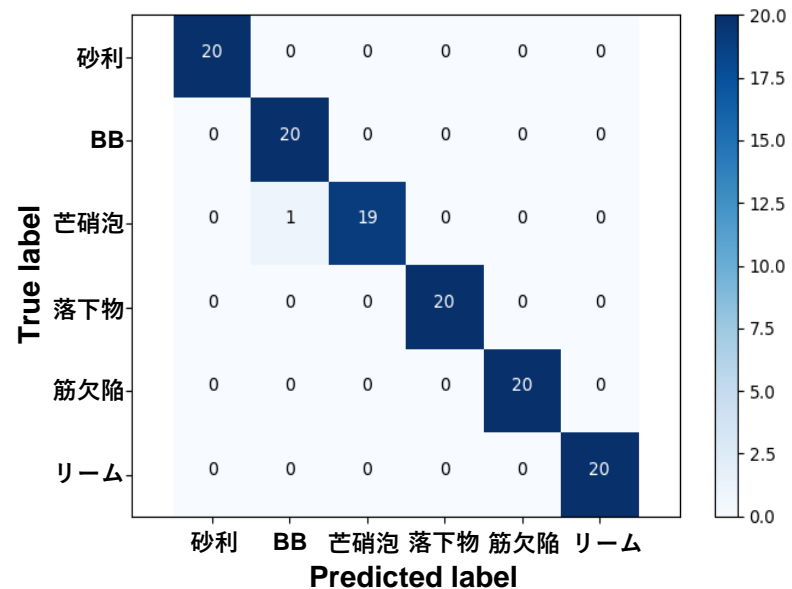


Multi-class metrics

● Accuracy ... 0.95

● Recall

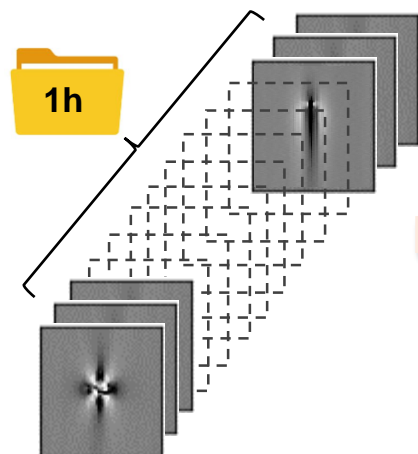
砂利	... 1.0	落下物	... 1.0
BB	... 1.0	筋欠陥	... 1.0
芒硝泡	... 0.95	リーム	... 1.0
		Ave.	... 0.99



製造工程への異物分類ツールの展開

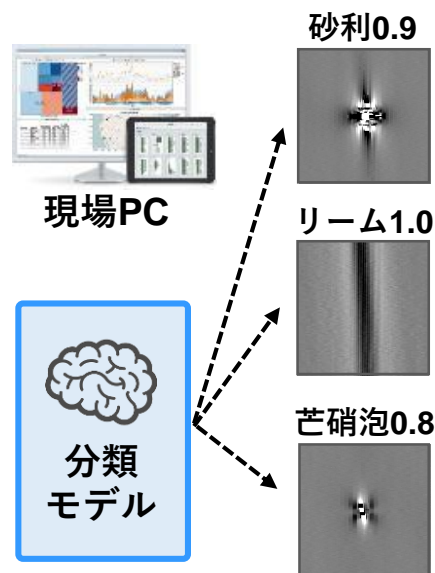
詳細な異物発生状況の可視化

1. データの準備

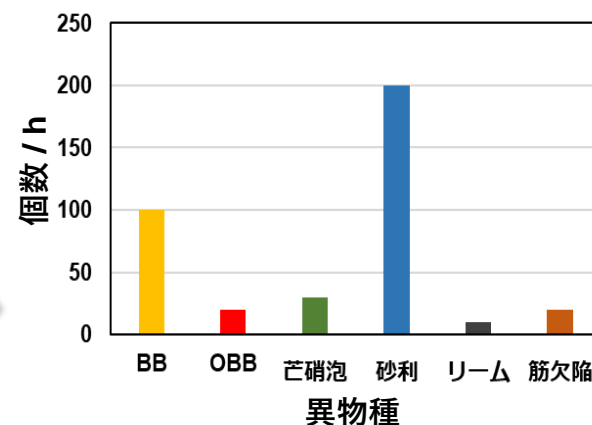


※外部PCで利用できるGUIは作成済
※モデルは随時更新する

2. 画像予測



3. 異物発生状況の可視化



4. 生産条件へフィードバック

【今後の予定】

- ①異物種の追加、品種別の予測モデルの作成
- ②フィードバック方法との紐づけ

データサイエンス部会の活動

7月末に硝子研の部会活動として立ち上げ (Python経験者6名で構成)



実務を通したデータサイエンスの土台作り

データの事前処理

データ形式
変数処理・欠損値処理

逆問題 (探索)

機能
特性

構造
組成

データの可視化

重要変数の抽出
相関関係・因果関係

順問題 (予測)

構造
組成

機能
特性

共通の課題

硝子研データベース (ガラス組成-熱膨張率等)

各チームの課題抽出

添加量の最適化、XRDピーク分類、etc...

研究開発に反映しやすい環境構築

データ収集形式の統一

	F	G	H	I	Q	R	S	T	U	V
1	SiO2	B2O3	Al2O3	MgO	TiO2	ZnO2	CeO2	P2O5	H2O	FractureT
2	67.76	0	2.33	8.68	0	1.12	0	0	0	0.74
3	62	0	13	3	0	2	0	0	0	0.89
4	71	0	1	0	0	16	0	0	0	0.7
5	60	0	22.5	0	0	0	0	0	0	0.69
6	55.15	0	14.38	0	0	0	0	0	0	0.81
7	60	0	0	0	0	0	0	0	0	0.91
8	67.2	10.6	10.6	2.9	0	0	0	0	0	0.9
9	57	19	0	0	0	0	0	0	0	0.68
10	66.42	0	4.76	3.44	0	1.69	0	0	0	0.65
11	61.9	0	13	3	1	0.5	0	0	0	0.94
12	66.2	7.6	11.3	5.3	0	0	0	0	0	0.9
13	85	0	0	0	0	0	0	0	0	0.8
14	66.7	0	0	0	0	0	0	0	0	0.74

解析ツールの導入検討



MATLAB



Minitab

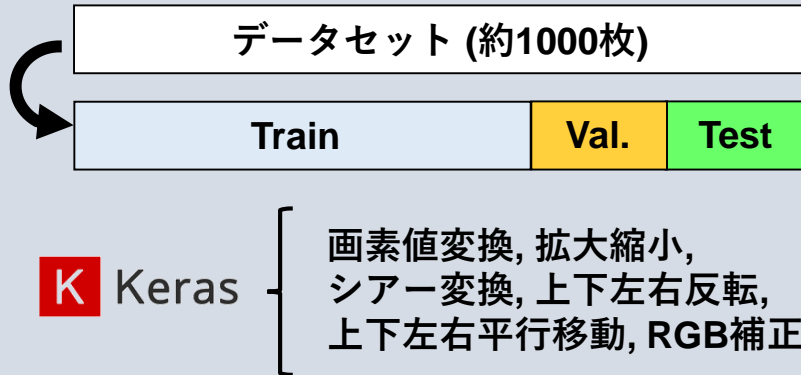
DataRobot



+ セミナー等を通したAIリテラシーの蓄積、硝子研内や化研との情報共有の場としての活用

画像分類モデル構築の流れ

1. Data split & augmentation



3. Hyperparameter tuning

Learning rate = $1e-3$, $1e-5$?
Batch size = 8, 32, 64 ?
Epoch = 10, 100, 1000 ?
Optimizer = Adam, RMSprop, SGD ?
Activation = Softmax, Sigmoid ?
or
全探索、ランダムサーチ、ベイズ最適化

2. Model building with Fine-tuning

