第9回FCワークシヨツプ MIX platform説明会

NIMS 袖山慶太郎, 冨中悟史

2024年10月2日 13:00-16:30

東京国際交流館 平成プラザ 3F メディアホール (オンライン配信および講習会会場)

Contents

前半：講演

1. MI研究の動向紹介（MIでできること、今後の動向）
2. MIX platform紹介
 - 2.1. 開発背景と現状、今後の利用について
 - 2.2. 電気化学測定結果の解析（LSV, IV, ORR）
 - 2.3. 簡単なMI解析
 - 2.4. 自動自律実験による実験データ蓄積
 - 2.5. 自然言語処理による論文データ蓄積
3. データの連携、検索

後半：実習(オンサイトのみ)

- ① MIX Platformへの接続方法、使い方の説明
- ② 演習課題の説明

MI解析の目的

データ科学、計算科学、理論及び実験を融合させ、物質・材料研究開発のあり方を革新し、**研究開発のスピードを大幅に加速させる**ことを目指す。

材料研究
(マテリアルリサーチ)

+

データ駆動科学、
機械学習

- ・ 超伝導材料 (無機固体、有機固体)
 - ・ 磁性材料 (無機固体)
 - ・ 高分子ポリマー材料 (有機液体、固体)
 - ・ 触媒材料(固体+気体、液体)
 - ・ リチウムイオン電池(無機固体+有機液体)
- など、材料研究の適用先は多岐にわたる

- ・ **新規材料開発の高速化**
- ・ **新規メカニズムの解明**

MI解析に必要な知識・ノウハウ

- 材料研究を加速するためにMI解析を利用する。

- **データのため方** + **解析の仕方** 両方の理解必要

自分のデータだけ使う

自分のグループのデータだけ使う

外部データと合わせて使う

- 語彙辞書の作成
- 計測ルールの制定

どんなアウトプットが欲しいか
どんなインプットがあるか
どんなMI解析手法が使えるか

- 要因解析
- 条件最適化
- 分類
- 新規材料提案

MI研究事例紹介

1. メタノール酸化触媒 活性の高い触媒組成を高速探索
初期17点 ベイズ最適化
2. Pt-HEA(111)モデル触媒の合成条件検討
活性の高い触媒組成と合成温度を高速探索
初期31点 ベイズ最適化
3. MI手法開発 分子式の方法のみから物性予測 SMILES-X

最近のMI研究の動向 (成功事例と課題 1)

メタノール酸化触媒の設計: 触媒活性の高いPtPdAu合金 の組成探索

Pt: Pd: Au = x: y: z

アクティブラーニングにより、
実験と予測を繰り返す。

Output: 活性の高い触媒の
組成を早く見つけたい
Input: 初期17点
MI手法: ベイズ最適化

最近のMI研究の動向（成功事例と課題1）

メタノール酸化触媒の設計: 触媒活性の高いPtPdAu合金 の組成探索

Pt: Pd: Au = x: y: z

アクティブラーニングにより、
実験と予測を繰り返す。

Pt: Pd: Au=59:40:1

の組成で従来より40%性能向上（予想外の組成が得られた）

Output: 活性の高い触媒の
組成を早く見つけたい
Input: 初期17点
MI手法: ベイズ最適化

・実験者にとって予想外の
組成が高速に見つかった。

最近のMI研究の動向 (成功事例と課題2)

Pt-HEA(111)モデル触媒の
合成条件検討
(東北大学 和田山先生)

ハイエントロピーアロイ
触媒の組成と実験条件を
アクティブラーニングに
より最適化。

Output: 高性能新規触媒組成と合成温度を提案したい
Input: 初期
MI手法: ベイズ最適化

初期
5k PCs後
活性

・触媒組成情報と合成温度のみから物性予測に成功

- ✓ 1st ~ 4th batchまでは、batch内で最高の特性を示した試料の”5kPCs後活性”がbatch毎に向上 ⇒ 4th batchで Pt / 等組成Cr-Mn-Fe-Co-Ni と同等のORR特性
- ✓ 非等組成比5元系, 4元系の合金化元素でも高ORR特性を示す合成条件を発見

最近のMI研究の動向 (成功事例と課題3)

分子式の情報のみから物性予測・新分子生成 (言語)

Input: SMILES strings (only)
Output: Physical properties prediction

Water solubility prediction
Interpretation is also possible.

(G. Lambard et al., *Mach. Learn.: Sci. Technol.*, 1(2), 025004, (2020).)

Output: 分子の物性を予測したい。高性能新規分子を提案したい
Input: 初期1783点, 106点
MI手法: SMILES-X

・分子式の情報のみから物性予測に成功した。

glass transition temperature (T_g)

Linear thermal expansion (β)

SMILES-X: using chemical structure information, not chemical fingerprint

Property	Model	R ² -score	RMSE
α (10^{-5} K^{-1})	SMILES-X	0.68 ± 0.06	2.44 ± 0.17
	Random forest	0.65 ± 0.02	2.54 ± 0.07
T_g (K)	SMILES-X	0.92 ± 0.00	33.21 ± 0.68
	Random forest	0.84 ± 0.00	45.70 ± 0.62

Only monomer information, β value is predicted.

E. Gracheva, G. Lambard, S. Samitsu, K. Sodeyama, A. Nakata, *STAM-methods*, 1:1, 213 (2021).

最近のMI研究の動向（今後の方針）

一般的な材料研究

データを使った材料探索（内挿）

新材料アイデアの
ひらめき

データを使った材料探索 (外挿)

- 大量のデータ蓄積により、これまで調べていない材料系を提案
- 新しいMI探索手法を提案

- ① PEFCデータベース
- ② 自律自動実験データ
- ③ 自然言語処理による論文データ
- ④ PEFCに関するMI解析ノウハウと解析手法開発

⇒ 新規材料の設計指針提案

最近のMI研究の動向（データのため方）

- 材料研究を加速するためにMI解析を利用する。
- **データのため方 + 解析の仕方 両方の理解必要**

自分のデータだけ使う

自分のグループのデータだけ使う

外部データと合わせて使う

項目名(語彙)や計測の仕方は仲間内だけで統一すればよかった。

しかし、それだと**データ数が少ない**。

データ数が多いほど、予測精度はあがり、新しい材料を見つけられる可能性高まる。

そこで、

- **語彙のリスト化（語彙の類義語を辞書化）**
- **計測ルールの共通化（理想的な計測範囲を設定）**

が、分野ごとに行われている。

MIX Platformについて (背景)

- NEDO燃料電池事業にてデータを蓄積するために開発
- ウェブアプリに機能を統合。自動的にDBにアクセス。
- データ蓄積時に、簡単なMI解析や実験解析を行うことが可能

MIX Platformについて（全体像）

MIX Platform: 初のPEFC専用データ解析・蓄積プラットフォーム

1. PEFC関連データを蓄積するだけでなくデータ解析が可能

- ・電気化学測定データ
- ・スペクトルデータ
- ・画像データ

本日は、現在開発が一番進んでいる2項目について説明します。
(今後アップデートしていきます。)

2. 蓄積したデータのMI解析機能

3. PEFC関連の生データだけでなく解析後のデータを、 実験条件などのメタデータ付きで蓄積が可能

MIX Platformで出来ること

ウェブブラウザ（Chrome推奨）でURLにアクセス

→データをアップロードする画面が表示
アップロード後、以下の機能が利用可能。

1. 生データの読込と実験データ解析
2. 統一的な形でデータを保管
→PEFCデータベースへ
3. 簡単なMI解析を手元で実施

エクスプローラー

ウェブアプリに
ドラッグ&ドロップ

MIX Platformで出来ること (データの読込と保存)

現在

20年後

ファイル形式
A

ファイル形式
C

ファイル形式
B

ファイル形式
D

変換
DB格納

HDF※形式

提供者のメリット①:
生データを自動読み込み
実験データ解析可能

提供者のメリット③:
簡単なMI解析ができる

※ 生データと実験条件
(メタデータ)を合わせて保存

読めない、、、ゴミ

提供者のメリット②:
データを死蔵させない

読める

設計利用

日本(NEDO事業)のメリット:
知の集約/ 材料設計に活用

MIX Platformで出来ること (DB)

- **PF マネジメントグループ(竹内)**
 - メンバーの追加/削除と権限付与
 - プロジェクトの作成・削除
 - データの閲覧・編集・消去

- **PF MIグループ (袖山)**
 - データの閲覧(編集は不可)

- **材料研究グループ (34事業のGL)**
 - 自分のPJのデータのみ閲覧可能
 - PJ内でデータの書き込み・消去
 - グループ内の利用者をリスト化→読み書きできるユーザーを管理者が追加
 - ×他のプロジェクトの閲覧

- ・異なるプロジェクトとのデータのコンタミはありません。
- ・自課題以外にデータにアクセスできるのは、現在袖山と竹内の二人だけです。
- ・運用は、基本的にMETI公開のガイドラインに沿って行います。

- ・PJ間のデータ(例えばK01とK02)を合わせてMI解析する場合があります。(事前に事業者の許可を取ります。)
- ・システム管理者(外部運用ベンダー)とNIMSの間で秘密保持事項を契約書に記載

MIX Platformで出来ること（電気化学解析）

- ・現在の電気化学測定結果の解析

1. 実験機器から(機器付属のソフトを用いて)生データを得る。

2. 測定者がExcelなどを用いて電気化学解析手法のプログラムを作る。

- × 測定者によりプログラム異なる。⇒出てくる値が異なる。

3. 測定者がExcelなどを用いて電気化学解析手法を実施する。

- × 自分でプログラムを書いていない場合、使い方を覚える必要がある。

MIX Platformで出来ること（電気化学解析）

- **MIX platformによる電気化学測定結果の解析**
 - 1. 実験機器から得られる生データを読み込み可能。**
 - データの描画実施。HDFファイルに自動変換
 - 現状は、フォーマットを限定。
 - 2. FC-Cubicのプロトコルでよく使われる解析コードを実装。**
 - 3. データ蓄積時にクリックするだけで解析可能。**

MIX Platformで出来ること（電気化学解析）

最新版のNEDO 評価解析プロトコルには以下のMEAの評価方法が記載されている。

(3)電極触媒・電極触媒層

A. 評価方法

・C-1	試験名: I-V測定方法・過電圧分離解析方法
・C-2 (III-3-4)	試験名: ECA測定方法
・C-3	試験名: クロスオーバー電流測定方法
・C-4 (III-3-2)	試験名: ORR活性評価方法
・C-5	試験名: HOR性能(アノード触媒)
・C-6	試験名: 酸素拡散抵抗
・C-7	試験名: プロトン輸送抵抗
・C-8	試験名: 高温評価(セル温度120℃まで)

電気化学的特性測定GrがMEA評価として標準的に実施している評価方法のうちよく利用される以下のものをMIX platformに実装した。

- ECSA評価 (CV測定)
- LSV測定の解析手法 (クロスオーバー電流(COC)測定方法)
- I/V測定測定の解析手法 (I-V測定方法・過電圧分離解析方法)
- ORR活性測定の解析手法 (ORR活性評価方法)

MIX Platformで出来ること（電気化学解析）

The screenshot shows the MIX Platform web interface. The browser address bar displays `https://v4.pefcdb.app/main`. The left sidebar contains a navigation menu with the following items: **ファイルブラウザ** (circled in red with a '1'), ファイルサイズ, 電気化学測定, プロット・解析, データの閲覧, MI解析, データ閲覧, データ修正, プロット (散布図), ヒストグラム (データ分布), 相関係数のヒートマップ, ベイズ最適化, 自然言語プロット, 線形回帰, and 非線形回帰. The main content area shows the current project as `workshop2410_public_project` and the current directory as `/演習用MEAデータ_最終版__1727422722727/raw_data/`. Above the file listing are buttons for `EDIT METADATA`, `UPDATE HDF`, **LOAD HDF** (circled in red with a '3'), `DELETE`, and `DOWNLOAD`. The file listing includes folders like `bayesian_optimization_sample`, `data1_solvents__1695788257042`, `JMCA_data__1695367776948`, `MIX講演会_演習__1695817909088`, `PCCP__1695805823753`, `Polymer_data__1695367796954`, `Pt_data__1695629475048`, and **演習用MEAデータ_最終版__1727422722727** (circled in red with a '2'). There are also options to create new folders.

- ① 「ファイルブラウザ」をクリック
- ② 「演習用MEAデータ」をクリック
- ③ 「LOAD HDF」をクリック

MIX Platformで出来ること（電気化学解析）

The screenshot displays the MIX Platform interface. On the left is a sidebar with navigation options: ファイルブラウザ, ファイルサイズ, 電気化学測定, プロット・解析, サイクリックボルタモグラム, リニアスイープボルタモグラム, 解析, データの閲覧, MI解析, データ閲覧, データ修正, プロット（散布図）, ヒストグラム（データ分布）, and 相関係数のヒートマップ. The main area shows the current project as 'workshop2410_public_project' and the current directory as '/演習用MEAデータ_最終版__1727422722727/raw_data/演習用MEAデータ_最終版/'. A file list on the left includes folders like 'bayesian_optimization_sample', 'data1_solvents__1695788257042', 'JMCA_data__1695367776948', 'MIX講演会_演習__1695817909088', 'PCCP__1695805823753', 'Polymer_data__1695367796954', 'Pt_data__1695629475048', and a folder '演習用MEAデータ_最終版__17274227227' which is highlighted with a red box and a circled '4'. To the right of the file list are buttons for 'LOAD', 'DELETE', and 'DOWNLOAD', with the 'DOWNLOAD' button highlighted by a red box and a circled '5'. Below the file list, a preview of the selected file '演習用MEAデータ_20240920_最終版.xlsx' is shown, including its type (xlsx), size (1.03 MB), and last modified date (24/09/27 16:39).

④ 「演習用MEAデータ」… 「演習用MEAデータ.xlsx」をクリック

⑤ 「DOWNLOAD」をクリック
⇒ 今日使うExcelファイルがローカルPCにダウンロードされる。

MIX Platformで出来ること (電気化学解析)

A	B	C	D	E	F	G
時間[sec]	電圧[V]	電流[A]	電流密度[mA/cm ²]	サイクル		
24.00259939	0.49965489	0.004702222	1.56740729	1		
24.00809939	0.49942899	0.001690158	0.563386072	1		
24.01809939	0.4989104	0.001724736	0.574911911	1		
24.02809939	0.49838904	0.001586235	0.52874507	1		
24.03809939	0.49790987	0.001563894	0.521297985	1		
24.04809939	0.49739963	0.001526072	0.508690642	1		
24.05809939	0.49690184	0.001556885	0.518961737	1		
24.06809939	0.49639705	0.001539803	0.513267728	1		
24.07809939	0.49591225	0.001525499	0.508499789	1		
24.08809939	0.49539247	0.001532964	0.510988042	1		
24.09809939	0.49490616	0.001517367	0.505788877	1		
24.10809939	0.49441218	0.001545508	0.515169288	1		
24.11809939	0.49389553	0.001554033	0.518010979	1		
24.12809939	0.49340263	0.001474073	0.491357757	1		
24.13809939	0.49289465	0.00147367	0.491223416	1		
24.14809939	0.49241871	0.001515829	0.505276371	1		
24.15809939	0.49189886	0.001537979	0.512659823	1		
24.16809939	0.49139088	0.001426772	0.475590612	1		
24.17809939	0.49089423	0.001517494	0.505831276	1		
24.18809939	0.49039876	0.001500783	0.500261005	1		
24.19809939	0.48991707	0.001439019	0.4796729	1		
24.20809939	0.48940638	0.001467807	0.489268907	1		
24.21809939	0.48891461	0.001530854	0.510284666	1		
24.22809939	0.48840522	0.001446403	0.489161142	1		

CV-metadata_40deg CV-data_40deg LSV-metadata_80deg | ...

相関係数のヒートマップ

722727/raw_data/演習用MEAデータ_最終版/

LOAD DELETE DOWNLOAD

演習用MEAデータ_20240920_最終版.xlsx

新規フォルダを作成...

演習用MEAデータ_20240920_最終版.xlsx

←実験データファイルの中身を確認

Sheet1: CV
Sheet3: LSV (COC)
Sheet5: ORR
Sheet7: IV

が記載されている。

- ④ 「演習用MEAデータ」… 「演習用MEAデータ.xlsx」をクリック
- ⑤ 「DOWNLOAD」をクリック
⇒ 今日使うExcelファイルがローカルPCにダウンロードされる。

MIX Platformで出来ること (電気化学解析)

MIX Platform

EXPORT PPTX

/演習用MEAデータ_最終版/演習用MEAデータ_20240920_最終版.xlsx / Sheet 1 (0000) ▼ to point

⑥ プロット・解析

⑦ サイクリックボルタモグラム

⑧ calculate area

CV

i (mA/cm²)

E (V vs. RHE)

— cycle 0 (lower) — cycle 0 (upper) — cycle 1 (lower) — cycle 1 (upper) — cycle 2 (lower)
— cycle 2 (upper) — cycle 3 (lower) — cycle 3 (upper) — cycle 4 (lower) — cycle 4 (upper)

Variables

⑦ X Label: 電圧[V] Y Label: 電流[A]

⑥ 「プロット・解析」、
「サイクリックボルタ
モグラム」をクリック

⑦ X Labelを「電圧」、
Y Labelを「電流」に
セットする。

⑧ Calculate areaを
オンにする。

MIX Platformで出来ること（電気化学解析）

MIX Platform

EXPORT PPTX

/演習用MEAデータ_最終版/演習用MEAデータ_20240920_最終版.xlsx / Sheet 1

calculate area

CV

CV

left endpoint
0.095449954

right endpoint
0.3589206

Extremum
0: x=0.500 y=37.438

COMPUTE AREA

Variables

X Label
電圧[V]

left endpoint
0.055004779

right endpoint
0.4399561

Extremum
0: x=0.440 y=60.223

Y Label
電流[A]

⑥ 「プロット・解析」、
「サイクリックボルタ
モグラム」をクリック

⑦ X Labelを「電圧」、
Y Labelを「電流」に
セットする。

⑧ Calculate areaを
オンにする。
⇒左のような画面にな
る。

⑨ 下の方にある
Parametersをセット
する。（次ページ）

MIX Platformで出来ること (電気化学解析)

データの閲覧

MI解析

データの閲覧

データの修正

プロット (散布図)

ヒストグラム (データ分布)

相関係数のヒートマップ

ベイズ最適化

自然言語プロット

線形回帰

非線形回帰

次元削減

マニュアル

書類作成

ログアウト

Extremum
0: x=0.500 y=37.438

COMPUTE AREA

Variables

X Label
電圧[V]

Extremum
0: x=0.440 y=60.223

Y Label
電流[A]

⑨ Parameters

データ取得間隔 [V/sec]
0.05
値は-1000~1000で0以外の実数値で入力してください。

単位表面積あたり吸脱着電気量 [$\mu\text{C}/\text{cm}^2$]
210
値は-1000~1000で0以外の実数値で入力してください。

白金担持量 [μg]
211
値は0~1000000で0以外の実数値で入力してください。

ノイズ除去

幾何学面積 [cm^2]
3
値は0~1000で0以外の実数値で入力してください。

⑩ APPLY

⑧ Calculate areaをオンにする。
⇒左のような画面になる。

⑨ Parametersを実験条件に合わせて入力する。(今回は左記のとおりの数値を入力)

⑩ 「APPLY」をクリック

MIX Platformで出来ること (電気化学解析)

MIX Platform

EXPORT PPTX

/演習用MEAデータ_最終版/演習用MEAデータ_20240920_最終版.xlsx / Sheet 1

calculate area

to point

⑪ COMPUTE AREA

⑫ Cycle 0

CV CV

left endpoint
0.095449954
right endpoint
0.3589206
Extremum
0: x=0.500 y=1.567
ECDSA = 67.28973

left endpoint
0.055004779
right endpoint
0.4399561
Extremum
0: x=0.440 y=2.521
ECDSA = 66.49959

Variables
X Label
電圧[V]
Parameters
Y Label
電流[A]

⑩ 「APPLY」をクリック

⑪ 「COMPUTE AREA」をクリック
⇒ ECDSAの値が計算される。

⑫ 「Cycle 0」を他のサイクルに変更して
COMPUTE AREAをクリックするとサイクルごとのECDSAが計算される

MIX Platformで出来ること (LSV測定COC解析)

目的

- 水素クロスオーバー電流 j_{xO} の見積もり
→ 電解質膜のガスバリア特性、膜劣化時の指標
- 短絡抵抗 R_{SC} の測定 → セル組立時に不備がないかなど

測定方法

評価解析プロトコル, C-3, p21

電位スイープ範囲・速度	<u>0.2~0.5V vs. RHE の範囲を低電位側から 0.5mV/s でスイープする。</u>
リーク電流の算出方法	0.2~0.5V の電位間をスイープしたときに得られる電位/電流密度曲線の <u>0.4~0.5V 近傍の直線部を 0V まで外挿したときの切片の電流密度が</u> クロスオーバー電流となり、傾きの逆数が電子抵抗(短絡抵抗)となる。

MIX Platformで出来ること (LSV測定COC解析)

MIX Platform

COC解析

解析情報入力

解析に用いるファイルを選択

① /演習用MEAデータ_最終版/演習用MEAデータ_20240920_最終版.xlsx / Sheet 3 (0000) ▼

解析に用いるデータの選択

計測時間 [sec] ② 時間[sec] ▼

電圧 [V] 電圧[V] ▼

電流 [A] 電流[A] ▼

計算に用いるパラメータの入力

セル面積[cm²]
3

③ 値は0~1000で0以外の実数値で入れてください。

EXECUTE ANALYSIS

①解析に用いるファイルで「Sheet3」を選択。

②解析に用いるデータの選択で、左記のように項目を選ぶ。（選択候補に挙がる項目は、Excelファイルの項目名が表示される）。

③「EXECUTE ANALYSIS」をクリック。

MIX Platformで出来ること (LSV測定COC解析)

- COC COC解析
- IV IV解析
- ORR ORR解析
- データの閲覧
- ★ MI解析
- データ閲覧
- データ修正
- プロット (散布図)
- ヒストグラム (データ分布)
- 相関係数のヒートマップ
- ベイズ最適化

電流 [A] 電流[A]

計算に用いるパラメータの入力

セル面積[cm²]
3

③ 値は0~1000で0以外の実数値で入力してください。

EXECUTE ANALYSIS

解析結果

算出値

④

短絡抵抗[kΩ cm ²]	31.242279576435834
水素クロスオーバー電流 [mA/cm ²]	2.6686489772355024

プロット用データ

EXPORT CSV

①解析に用いるファイルで「Sheet3」を選択。

②解析に用いるデータの選択で、左記のように項目を選ぶ。(選択候補に挙がる項目は、Excelファイルの項目名が表示される)。

③「EXECUTE ANALYSIS」をクリック。

④短絡抵抗と水素クロスオーバー電流の値が計算される。

MIX Platformで出来ること (IV解析)

- 目的**
- MEAのIV特性評価→触媒、電解質膜、GDLなどの全構成要素の正味のIV性能
 - 過電圧分離→触媒活性、電解質膜抵抗、酸素輸送抵抗の要素を分離

測定方法 評価解析プロトコル, C-1, p19

	1cm×1cmセル	1cm×3cmセル	JARI標準セル
電流密度と保持時間	①** 2,4,6,10,20,50,75,100 mA/cm ² (各5min)	① 2,4,6,10,20,50,75,100 mA/cm ² (各5min)	① 25,50,75,100 mA/cm ² (各5min)
低電流密度から 高電流密度へシフト	② 200,400,600,800,1000, 1500,2000mA/cm ² (各10min)	② 200,400,600,800,1000, 1200 mA/cm ² (各10min)	② 200,400,600,800,1000, 1200 mA/cm ² (各10min)

←各電流密度での定常状態の電圧を計測

MIX Platformで出来ること (IV解析)

The screenshot shows the MIX Platform interface for IV analysis. The left sidebar contains navigation options: ファイルブラウザ, ファイルサイズ, 電気化学測定, プロット・解析, サイクリックボルタモグラム, リニアスイープボルタモグラム, 解析, COC解析, IV解析, ORR解析, データの閲覧, MI解析, and データ閲覧. The main content area is titled 'MIX Platform' and 'IV解析'. Under '解析情報入力', there are three sections: '解析に用いるファイルを選択' (1), '解析に用いるデータの選択' (2), and '計算に用いるパラメータの入力' (3). The first section shows a file selection dropdown with the path '/演習用MEAデータ_最終版/演習用MEAデータ_20240920_最終版.xlsx / Sheet 7 (0000)'. The second section shows three dropdown menus for '電圧 [V]', '電流 [A]', and '測定抵抗 [mΩ]'. The third section shows input fields for 'セル面積 [cm^2]' (value: 3) and '理想起電圧 [V]' (value: 1.17). A green button labeled 'EXPORT ANALYSIS RESULT CSV' is at the bottom, annotated with (4).

①解析に用いるファイルで「Sheet7」を選択。

②解析に用いるデータの選択で、左記のように項目を選ぶ。（選択候補に挙がる項目は、Excelファイルの項目名が表示される）。

③計算に用いるパラメータを左記のように入力。

④「EXPORT ANALYSIS RESULT CSV」をクリック。

MIX Platformで出来ること (ORR解析)

目的

- 触媒性能の指標となる質量活性と面積比活性を算出する。

測定方法

評価解析プロトコル, C-4, p22~23

	1cm×1cmセル	JARI標準セル
電流密度と保持時間	① 0 mA/cm ² (O ₂ 導入後2min) ② 1000, 200, 100, 80, 60, 40, 20 mA/cm ² (各15min)	① 0 mA/cm ² (O ₂ 導入後2min) ② 1000, 200, 100, 80, 60, 40, 20 mA/cm ² (各15min)
高電流密度から 低電流密度へシフト	必要に応じて ③ 10, 6, 4, 2 mA/cm ² (各5min)	

MIX Platformで出来ること (ORR解析)

The screenshot shows the 'ORR解析' (ORR Analysis) page in the MIX Platform. The interface includes a sidebar with navigation options and a main content area for data input. Four red callouts are present:

- ①** A dropdown menu for selecting the file to use for analysis, showing the path: `/演習用MEAデータ_最終版/演習用MEAデータ_20240920_最終版.xlsx / Sheet 5 (0000)`.
- ②** A dropdown menu for selecting the data to use for analysis, with options: 電圧[V], 電流[A], and 測定抵抗[Ω cm²].
- ③** A form for entering calculation parameters, including:
 - セル面積[cm²]: 3
 - 水素クロスオーバー電流[mA/cm²]: 2.71
 - 白金担持量[mg/cm²]: 0.211
 - ECSA[m²/g]: 50.96
- ④** A green button labeled 'EXECUTE ANALYSIS' at the bottom of the form.

①解析に用いるファイルで「Sheet5」を選択。

②解析に用いるデータの選択で、左記のように項目を選ぶ。（選択候補に挙がる項目は、Excelファイルの項目名が表示される）。

③計算に用いるパラメータを左記のように入力。

④「EXECUTE ANALYSIS」をクリック。

MIX Platformで出来ること (ORR解析)

- COC解析
- IV解析
- ORR解析
- データの閲覧
- ★ MI解析
- データ閲覧
- データ修正
- プロット (散布図)
- ヒストグラム (データ分布)
- 相関係数のヒートマップ
- ベイズ最適化
- 自然言語プロット
- 線形回帰
- 非線形回帰

計算に用いるパラメータの入力

セル面積[cm²]

3

値は0~1000で0以外の実数値で入力してください。

水素クロスオーバー電流[$\mu\text{A}/\text{cm}^2$]

2.71

値は-1000~1000の実数値で入力してください。

白金担持量[mg/cm^2]

0.211

値は0~1000で0以外の実数値で入力してください。

ECSA[m^2/g]

50.96

値は0~1000で0以外の実数値で入力してください。

EXECUTE ANALYSIS

解析結果

算出値

⑤ 質量活性[A/g]	109.17864786067081
面積比活性[$\mu\text{A}/\text{cm}^2\text{-Pt}$]	214.24381448326298

プロット用データ

EXPORT CSV

①解析に用いるファイルで「Sheet5」を選択。

②解析に用いるデータの選択で、左記のように項目を選ぶ。(選択候補に挙がる項目は、Excelファイルの項目名が表示される)。

③計算に用いるパラメータを左記のように入力。

④「EXECUTE ANALYSIS」をクリック。

⑤質量活性と面積比活性の値が計算される。

MIX Platformで出来ること (MI解析)

1. 表形式のデータを読み込み (.csvファイル)

➤既にきれいな形に揃っているデータを想定。

2. データのMI解析が可能

➤散布図、ヒストグラム分布、相関係数のヒートマップをウェブアプリ上で描画可能。

➤線形回帰 (重回帰、LASSO、Ridge回帰)を利用可能。

欲しい物性値の予測と、何が予測に効くかの要因解析をウェブアプリ上で実施。

1. 非線形回帰手法(**決定木**、**Random Forest**、**XGBoost**)の利用
2. **全状態探索** (線形回帰) 手法の利用
3. **ベイズ最適化**の利用
4. 主成分分析、UMAPの利用
5. その他 (データ上書き機能、PPTファイル出力機能...)

MIX Platformでできること (DBへの蓄積 1)

①ウェブブラウザから
アクセス可能
(windowsやMacで
利用可能)

②Drag & dropで
データを簡単に
アップロードできる
(テキストやエクセル
ファイルに対応)

↓
自動的にファイルの中身
を読み取って構造化

MIX Platformでできること (DBへの蓄積 2)

PEFC Data Viewer

Current Directory

sample_data_csv

新規フォルダを作成

Metadata entry

データ提供者: 山田 太郎

所属: NIMS

実験方法: CV

データ取得日: 2021/10/15

サンプル識別子 (任意): 白金

キーワード・タグ (任意): #うまくいった, 論文

実験状況 (任意)

項目名	値	
温度	26°C	🗑️
湿度	40%	🗑️

+ 項目を追加

備考

OK CANCEL

③ 必要最小限の実験条件 (メタデータ) をデータアップロード時に入力
↓
生データと共にDBに保存

④ データ提供者が保存しておきたい項目やメモは自由記述で保存可能 (キーワードはハッシュタグで保存)

MIX Platformでできること (DBへの蓄積 3)

The screenshot displays the MIX Platform's file management interface. On the left, a sidebar shows a list of folders and files, with 'data1_solvents__1695788257042' selected. The main area shows the details for 'data1_solvents.csv', including its type (csv), size (5.89 KB), and last modified date (23/09/27 13:18). At the top right, there are three buttons: 'LOAD CSV' (green), 'DELETE' (red), and 'DOWNLOAD' (green). A blue arrow points from the 'DOWNLOAD' button to a callout box. Another blue arrow points from a callout box to the selected file in the sidebar.

⑤ MacのFinder風に
ファイルを表示
(構造化データ + 生データ)

⑥ DOWNLOAD: ファイルを
自由にダウンロード可能
LOAD CSV: MIX Platformの
MI解析機能を利用可能

MIX Platformでできること (MI解析1)

CSVファイルの読込(MI解析はcsvファイルのみ対応)

The screenshot displays the MIX Platform interface. On the left is a sidebar with navigation options: ファイルブラウザ, ファイルサイズ, 電気化学測定, プロット・解析, データの閲覧, MI解析, データ閲覧, データ修正, プロット (散布図), ヒストグラム (データ分布), 相関係数のヒートマップ, ベイズ最適化, 線形回帰, 重回帰分析, Lasso回帰, and Ridge回帰. The main area shows the 'Current Project' as 'workshop2_public' and the 'Current Directory' as '/data1_solvents__1695788257042/raw_data/'. A file browser lists several folders, with 'data1_solvents__1695788257042' selected. A file named 'data1_solvents.csv' is highlighted. A dialog box asks: 'data1_solvents.csvを分析用データとしてロードしますか?' (Do you want to load data1_solvents.csv as analysis data?). Buttons for 'OK' and 'CANCEL' are visible. In the top right, there are buttons for 'LOAD CSV', 'DELETE', and 'DOWNLOAD'. A callout box with an arrow points to the directory path, containing the text: 'アップロードしたCSVファイルをLOAD' (Load the uploaded CSV file).

MIX Platformでできること (MI解析2)

Data Viewer (アップロードしたデータの確認、特徴量の選択/非選択)

①MI解析時に利用
するかどうかを
データごとに選択可能

The screenshot displays the MIX Platform interface. On the left is a sidebar menu with various options. The main area is titled 'Data Viewer' and shows a table with columns for 'Enabled', 'Name', and various chemical features. A blue callout box points to the 'Enabled' column, indicating that users can select or deselect features for MI analysis. A red box highlights the 'MI解析' option in the sidebar menu, with the text 'MI解析パート' next to it.

Enabled	Name	Flashpoint	Molecularweight	density	BoilingPoint	MeltingPoint	LUMO	HOMO	Dipolemoment	DELisolv
☒	1	204	434.63	0.93	215	-70	1.103	-7.622	3.387	-0.110336759
☒	2	178	218.33	1.22	310	59	-1.108	-6.111	1.597	-0.067238919
☒	3	170	119.12	1.36	240	98	-1.325	-6.694	3.792	-0.090462811
☒	4	165	120.17	1.26	285	28	0.826	-7.383	5.087	-0.101567304
☒	5	161	136.19	0.93	172	-3.5	0.497	-8.789	0.001	-0.07532871

Rows per page: 5 1-5 of 82

Dense padding

MIX Platformでできること (MI解析3)

散布図の描画(Scatter)

The screenshot displays the MIX Platform interface. On the left, a sidebar menu lists various functions: ファイルブラウザ, ファイルサイズ, 電気化学測定, プロット・解析, データの閲覧, MI解析 (highlighted with a red box and labeled 'MI解析パート'), データ閲覧, データ修正, プロット (散布図), ヒストグラム (データ分布), 相関係数のヒートマップ, ベイズ最適化, and 線形回帰. The main area shows a 'Scatter Plot' with 'MullikenO' on the x-axis and 'DELisolv' on the y-axis. Below the plot, dropdown menus for 'MullikenO' and 'DELisolv' are visible. A callout box with arrows pointing to these dropdowns contains the text: 'x 軸、y 軸を選択 (y軸が回帰分析時の目的変数になる)'.

MIX Platformでできること (MI解析4)

ヒストグラム

MI解析パート

アップロードデータの分布を描画

MIX Platformでできること (MI解析5)

相関係数のヒートマップ(Pearsonの相関係数)

MIX Platformでできること (MI解析5)

相関係数のヒートマップ(Pearsonの相関係数)

MIX Platformでできること (MI解析6)

線形回帰分析(重回帰分析)

MIX Platform

EXPORT PPTX

EXPORT COEFFICIENTS

Linear Regression of DELIsolv

True value vs Prediction (DELIsolv)

$R_f^2: 0.482$ $F_{cv}: 66.523$ $RMSE: 0.013$ $MAE: 0.097$

K-FOLD CROSS VALIDATION

k: 5

$R_f^2: \text{NaN} \pm 0.000$ $F_{cv}: \text{NaN} \pm 0.000$ $RMSE: \text{NaN} \pm 0.000$ $MAE: \text{NaN} \pm 0.000$

線形回帰
パート

S 重回帰分析

L Lasso回帰

R Ridge回帰

特徴量抽出
(重要な特徴量を描画)

アップロードした実験値と
回帰による予測値を比較

MIX Platformでできること (MI解析7)

線形回帰分析(LASSO分析)

MIX Platform

EXPORT PPTX

EXPORT COEFFICIENTS

Lasso Regression of DELisolv

Weight

Flashpoint Molecular weight BoilingPoint MeltingPoint LUMO HOMO Dipolemoment R LjO MullikenO

True value vs Prediction (DELisolv)

Prediction

True value

$R_f^2: 0.477$

RMSE: 0.013

MAE: 0.096

RELOAD

0.0 100.0

K-FOLD CROSS VALIDATION

k 5

$R_f^2: \text{NaN} \pm 0.000$

$F_{cv}: \text{NaN} \pm 0.000$

RMSE: NaN $\pm$ 0.000

MAE: NaN $\pm$ 0.000

線形回帰

重回帰分析

Lasso回帰

Ridge回帰

ハイパーパラメータの手動選択

MIX Platformでできること (MI解析7)

線形回帰分析(LASSO分析)

LASSO (least absolute shrinkage and selection operator)

$$E = \sum_{\mu=1}^{103} (z^{\mu} - \sum_{i=1}^{10} w_i x_i^{\mu})^2 + \lambda \sum_{i=1}^{10} |w_i|$$

- パラメータ λ を変化させることで予測精度を上げる

MIX Platformでできること (MI解析7)

線形回帰分析(LASSO分析) クロスバリデーション

K-FOLD CROSS VALIDATION

k

5

R^2_j : -3.399 ± 3.350

F_{cv} : 60.661 ± 107.605

$RMSE$: 48.381 ± 22.452

MAE : 5.620 ± 0.990

Cross Validation (CV)

1 10 20 30 40 50 60 70 80 90 100

■ : 回帰用 (学習用) データ

■ : 誤差確認用 (テスト用) データ

100点あるデータを全て使って予測モデルを作成
→新しいデータの予測性能がわからない。

100点データがあるうちの、90点を使って予測モデル作成。
残り10点を使って予測モデルにあてはめ、精度検証。

どの10点を残すかを変えながら10パターン計算を行う。

MIX Platformでできること (MI解析7)

線形回帰分析(LASSO分析) ハイパーパラメータ最適化

PARAMETER OPTIMIZATION

α min: 0 α max: 100 α count: 100

number of α values to try

k: 5 Holdout ratio: 0.2 Target loss: RMSE

cross validation

Holdout

R^2_f : -3.520

F_{cv} : -41.628

RMSE: 55.849

MAE: 6.452

Best cross-validation error

R^2_f : -3.409 $\pm$ 1.095

F_{cv} : 7.290 $\pm$ 78.316

RMSE: 28.590 $\pm$ 7.313

MAE: 4.555 $\pm$ 0.331

- パラメータ λ を変化させることで予測精度を上げる
- λ をクロスバリデーションによるCV errorの最小値をとる λ として選ぶ

MIX Platformでできること (MI解析8)

線形回帰分析(全状態探索)

線形回帰
パート

- S 重回帰分析
- L Lasso回帰
- R Ridge回帰
- E 全状態探索による線形回帰

計算に数十秒かかります。
余り沢山は実行しないで
頂けると助かります。

全ての説明変数の組合せ（10変数あれば1023通り）
に対して線形回帰を実施。予測精度の高いものから
順にどの説明変数を使ったか表示。

MIX Platformでできること (MI解析9)

非線形回帰分析(決定木)

- データ閲覧
- データ修正
- プロット (散布図)
- ヒストグラム (データ分布)
- 相関係数のヒートマップ
- ベイズ最適化
- 線形回帰
- 非線形回帰 **非線形回帰**
 - 決定木**
 - Random Forest
 - XGBoost
- 次元削減
- マニュアル
- ログアウト

MIX Platformでできること (MI解析10)

非線形回帰分析(Random Forest)

- データ閲覧
- データ修正
- プロット (散布図)
- ヒストグラム (データ分布)
- 相関係数のヒートマップ
- ベイズ最適化
- 線形回帰 非線形回帰
- 非線形回帰 **パート**
- 決定木
- Random Forest**
- XGBoost

MIX Platformでできること (MI解析11)

非線形回帰分析(XGBoost)

- データ閲覧
- データ修正
- プロット (散布図)
- ヒストグラム (データ分布)
- 相関係数のヒートマップ
- ベイズ最適化
- 線形回帰
- 非線形回帰** **非線形回帰**
パート
- 決定木
- Random Forest
- XGBoost
- 次元削減

MIX Platformでできること (MI解析12)

ベイズ最適化

例：2次元空間における最大値を探す

①. 初期値

x	y	z
-6	-8	3.07E-23
2	-1	0.013064
-3	2	0.0002
-10	5	1.14E-28
5	9	1.53E-24

②. 次の実験すべき点の提案

MIX Platform

NEXT EXPERIMENT

Bayes Optimization

Target

Enabled	Name	x	y	z
☒	Suggestion1	1.888	-0.9462	
☒	1	-6.000	-8.000	3.070e-23
☒	2	2.000	-1.000	0.01306
☒	3	-3.000	2.000	0.0002000
☒	4	-10.00	5.000	1.140e-28

Rows per page: 5 1-5 of 6

Dense padding

MIX Platformでできること (MI解析12)

ベイズ最適化

例：2次元空間における最大値を探す

①. 初期値

x	y	z
-6	-8	3.07E-23
2	-1	0.013064
-3	2	0.0002
-10	5	1.14E-28
5	9	1.53E-24

②. 次の実験すべき点の提案

Enabled	Name	x	y	z
☒	Suggestion1	1.888	-0.9462

③. 実験実施

(今回は実験の代わりにExcelで値を導出)

	A	B	C	D
1	試行回数	予測値x	予測値y	予測値z
2	0	-3	2	0.00023928
3	0	-10	5	1.14397E-28
4	0	5	9	1.52831E-24
5	0	-6	-8	3.0697E-23
6	0	2	-1	0.013064233
7	1	1.94	-0.9594	0.015300018
8	2			

MIX Platformでできること (MI解析12)

ベイズ最適化

例：2次元空間における最大値を探す

①. 初期値

x	y	z
-6	-8	3.07E-23
2	-1	0.013064
-3	2	0.0002
-10	5	1.14E-28
5	9	1.53E-24

②. 次の実験すべき点の提案

Enabled	Name	x	y	z
☒	Suggestion1	1.888	-0.9462	

MIX Platform

NEXT EXPERIMENT

Enabled	Name	x	y	z
☒	Suggestion1	1.888	-0.9462	

Target

Enabled	Name	x	y	z
☒	Suggestion1	1.888	-0.9462	0.015300018
☒	1	-6.000	-8.000	3.070e-23
☒	2	2.000	-1.000	0.01306
☒	3	-3.000	2.000	0.0002000
☒	4	-10.00	5.000	1.140e-28

Rows per page: 5 1-5 of 6

③. 実験実施

(今回は実験の代わりにExcelで値を導出)

④. 実験値をMIXに代入

7	1	1.94	-0.9594	0.015300018
8	2			

MIX Platformでできること (MI解析12)

ベイズ最適化

例：2次元空間における最大値を探す

①. 初期値

x	y	z
-6	-8	3.07E-23
2	-1	0.013064
-3	2	0.0002
-10	5	1.14E-28
5	9	1.53E-24

②. 次の実験すべき点の提案

Enabled	Name	x	y	z
☒	Suggestion1	1.888	-0.9462

③. 実験実施

(今回は実験の代わりにExcelで値を導出)

④. 実験値をMIXに代入

次の実験提案

Enabled	Name	x	y	z
☒	Suggestion2	1.435	-1.162	0.015300018
☒	Suggestion1	1.888	-0.9462	0.01530
☒	1	-6.000	-8.000	3.070e-23
☒	2	2.000	-1.000	0.01306
☒	3	-3.000	2.000	0.0002000

7	1	1.94	-0.9594	0.015300018
8	2			

MIX Platformでできること (MI解析13)

主成分分析

MIX Platformでできること (MI解析14)

UMAP

- プロット・解析
- データの閲覧
- ★ MI解析
- データ閲覧
- データ修正
- プロット (散布図)
- ヒストグラム (データ分布)
- 相関係数のヒートマップ
- ベイズ最適化
- 線形回帰
- 非線形回帰
- 次元削減
 - 主成分分析
 - UMAP

- ① PEFCデータベース
- ② 自律自動実験データ
- ③ 自然言語処理による論文データ
- ④ PEFCに関するMI解析ノウハウと解析手法開発

⇒ 新規材料の設計指針提案

自律・自動実験について： NIMSの自律自動実験装置の紹介

国立研究開発法人物質材料研究機構（NIMS）

マテリアル基盤センター

冨中 悟史、中田 謙吾

過去の講演内容

2022年 燃料電池材料の自律実験システムの設計

- ✓MIグループとしての自律実験
- ✓予測器のためのデータ処理について
- ✓自律実験のシステム設計の解説

2023年 燃料電池材料の自律実験システムの紹介

- ✓背景
- ✓NIMSのシステム（ELVES）の紹介
- ✓現状紹介（課題共有）

本日の講演内容

「燃料電池材料の自律実験システムの紹介」

- ✓全体のおさらい
- ✓動作の紹介

MIグループとしての自律実験の意義

- 均質かつ大量のデータ（データベース強化）
→MIによる材料探索が加速
- 人が記録しきれないプロセス条件などもデータ化
- 並列化が容易
- 属人性を排除（特に燃料電池触媒では大きい）

国全体の材料探索の効率化・迅速化

研究開発のDXの必要性

- ✓ 社会的要求に基づく研究開発は競争を伴う
- ✓ 既存の研究開発のやり方を変革することで**加速**
- ✓ 基礎研究 ≠ 役に立たないもの（DXで加速すべき研究とは何かを明確に）

燃料電池事業全体への加速効果

MI実施の視点でのメリット（前頁）に加え、材料探索スピードが少なく見積もって100倍加速の効果がある

加速アイテムの位置づけ

- 担体にナノ粒子を担持するという現行触媒材料のアーキテクチャの範囲で実材料となりえる手前までを高速で探索することを目標に設計。
- ナノ粒子として種々の合金を合成し、実作動に近い特性評価と光電子分光（化学状態や価電子情報の収集）により特性データと科学的な解析を行う。
- 産業界側での合成の最適化からは別のDX技術が必要だが、その手前までは広い範囲をカバー。

構想図：膨大な数のパラメータ空間を自動探索

ELVES：ロボットと人工知能の集合

ELVES 開発プロジェクト

「無人で研究を進める小人たち」

- 制御はPythonで自動化
- APIが無いものはNIMSで作製し公開
- データベース化（データプラットフォーム）
- 高速解析しMI解析（NIMSの専門家の知）
- AIをNIMSで構築し育成

統一規格品でなくても、インターフェースを作って連携させる。

目的：研究のボトルネックを解消するために、実験の並列処理と異なる機器の直列化を実現する開発。

ELVESの主要要素

各要素のプログラムを開発し、柔軟に高速処理を行う

- サーバー

真空クラスターシステム

- 真空クラスターシステム
- アークプラズマ装置
- 雰囲気制御光電子分光装置

データベース

燃料電池評価システム

- 電気化学測定システム
- セル
- サーボプレス
- ガス制御
- 搬送ロボット

①アークプラズマによる白金の堆積

真空システム

②光電子分光測定による分析

真空システム

Pt XPS 2d image

生データを記録しながら、高速処理で必要な情報を取り出す

```
#####
4
5 # Description: XPS data extractor for LIVES
6 # Version: 1.0.0
7 # Last update: 2023.3.22
8 # Developer: Satoshi Tomioka
9 # Affiliation: National Institute for Materials Science
10 # Contact: TOMIYAMA.Satoshi@nims.go.jp
11 #####
12 from __future__ import annotations
13 import copy, glob, os, re
14 #import numpy as np
15 #import arpes
16 #from arpes.io import load_data
17
18 #import peak
19 import hspy
20 from matplotlib import pyplot as plt
21
22
23
24 file_name = "h5_files/0231 test Pt_3Hz_100.h5"
25 data_folder = os.path.dirname(os.path.abspath(__file__))
26 print(data_folder)
27
28 def find_files():
29     files = glob.glob(data_folder+"/*"+file_name, recursive=True)
30     return files
31
32 class XPS:
33     def __init__(self):
34         self.data_file = None #hdfs
35
36     def read(self):
37         self.file = hspy.File(self.data_file, "r")
38
```


現状

- Pt触媒、PtCo触媒での試験中
- 日々、課題を1つずつ解決
- 類似の検討を進める方に向けた課題の取りまとめ

ありがとうございました。

謝辞：

この検討は、NEDOの「マテリアルズ・インフォマティクスによる燃料電池材料の研究開発」として実施して得られたものです。

文献情報抽出技術による持続的なデータベース構築

SC-CoMics: A Superconductivity Corpus for Materials Informatics

Kyosuke Yamaguchi¹, Ryoji Asahi², Yutaka Sasaki¹

¹Toyota Technological Institute

2-12-1 Hisakata, Tenpaku-ku, Nagoya, 468-8511 Japan

²Toyota Central R&D Labs., Inc., 41-1, Yokomichi, Nagakute, Aichi 480-1192, Japan

文献、特許から組成—プロセス—物性をデータベース化

Yamaguchi, Asahi, Sasaki, Adv. Eng. Inf. 54, 101768 (2022)

Example of the corpus construction

Document collection

1,000 abstracts
on superconductivity

Annotation design

Named Entity

Characterization,
Process, Property,
Material, Element,
Value, SC

Main Material

Event

Doping
Dopant
Site

Relation

Equivalent
Target
Condition

Manual annotation

Two workers

Tagged corpus

Annotation examples

- 1 In the history of superconductors, Fe-based superconductor LnFePnO_{1-x}F_x (1111) has the highest T_c at ~58K except Cu-based oxide superconductors.
- 2 Electron doping by substitution of O²⁻ with F⁻ was essential to induce superconductivity.
- 3 Local structures around iron are found to have a good empirical relation to T_c.
- 4 Maximum T_c is observed for 1111 materials having nearly regular tetrahedral FeAs₄ local structures.
- 5 The T_c and upper critical field of 1111 compounds are very attractive for power cable under high magnetic fields, although fabrication of practical superconducting wires is a tough challenge due to the difficulty in controlling the F contents during a processing for ex-situ powder-in-tube method.

brat rapid annotation tool: <https://brat.nlplab.org/index.html>

Extractions of superconducting materials and Tc

- ✓ Construct a new corpus, SC-CoMlcs, for superconducting material information extraction and apply deep learning to confirm the usefulness of the corpus in the superconducting materials search.

Extraction

Material **Trans.Temp.** **Doping**

Element	Tc	Dopant	Site
YBCO	-	Bi ₂ O ₃	-
Bi ₂ Sr ₂ CaCu ₂ O ₈	60K	-	-
La _{1.5} Ba _{1.5} Cu ₃ O ₇	-	Ca	La
NdBaSrCu ₃ O _{7-δ}	77K	Ca	Nd
⋮			

Extraction of the set of main material, phase transition temperature, and doping information

➔ **Prediction of characteristics of candidate materials
(Semi-)automatics population of materials databases**

Extraction of superconducting transition temperature (T_c)

extracted data

comparison with existing database (Supercon*)

- Each peak corresponds to representative superconducting materials
- Automatically extracted data is consistent with the existing database Supercon* that has been manually constructed.

Performance of the Tc prediction model

- The model can predict Tc within an error of 14 K only using automatically extracted data.

データ連携と検索

異なる研究者、分野のデータ連携→合体した大きな表を作成

自然言語処理による論文データ...

サンプル名

		特徴量			XRD	DFT	自動実験
サンプル名	事業者A データ	① 共通するデータを抽出					
	事業者B データ	② 共通する計測・計算データを利用					
		...					

1. 本事業の全てのデータを融合させ、一枚のテーブルを作る。
2. 個別データ間の関連性を考慮していないため、非対角成分にデータが無く、データ解析ができない。
3. 非対角成分にデータを与えるため、
 - ① 個別データ間で共通するデータを抽出する。(データの共通リスト化)
 - ② 個別データ間で共通する計測手法や計算データを利用する。

⇒大量に蓄積される、異なる性質のデータを利用した解析の実施が可能に。

事業の全データを融合させたテーブル

⇒MIX platformに必要なデータを検索し、共通するデータを自動連携させてMI解析を実行する機能を搭載(中)

データ連携と検索

**データ連携、検索パートは、まだ本番環境には実装されておらず、
以後デモバージョンの紹介となります。**

データ連携と検索

- ① ファイルブラウザ
- 検索
- ファイルサイズ
- 電気化学測定
- プロット・解析
- データの閲覧
- MI解析
- データ閲覧
- データ修正
- プロット (散布図)
- ヒストグラム (データ分布)
- 相関係数のヒートマップ
- ベイズ最適化
- 自然言語プロット
- 線形回帰
- 非線形回帰
- 次元削減
- マニュアル
- 書類作成
- ログアウト

Current Project : workshop2410_search_demo_project

Current Directory : /講習会用検索デモ__1727250072159/raw_data/講習会用検索デモ/

< >

LOAD DELETE DOWNLOAD

【注意】 このプロジェクトでファイル追加・移動・差し替え・削除は行わないでください！.txt	講習会用検索デモ	solvent_carbonate_mpbp.csv
講習会用検索デモ__1727250072159	新規フォルダを作成...	solvent_web1_mpbp.csv
新規フォルダを作成...		solvent_ether_mpbp.csv
		solvent_lactone_mpbp.csv
		solvent_nitrile_mpbp.csv
		新規フォルダを作成...

① 「検索」 を
クリック

データ連携と検索

MIX Platform

① ファイルブラウザ

検索

ファイルサイズ

電気化学測定

プロット・解析

データの閲覧

MI解析

データ閲覧

データ修正

プロット (散布図)

ヒストグラム (データ分布)

相関係数のヒートマップ

ベイズ最適化

自然言語プロット

線形回帰

非線形回帰

次元削減

マニュアル

書類作成

ログアウト

キーワード検索 詳細検索

検索キーワード一覧をカンマ区切りで入力してください

②

検索

- ① 「検索」をクリック
- ② にキーワードを入力
(入力しようとすると検索候補が自動表示)
⇒例えば melting point を選択

データ連携と検索

MIX Platform

① ファイルブラウザ

検索

ファイルサイズ

電気化学測定

プロット・解析

データの閲覧

MI解析

データ閲覧

データ修正

プロット (散布図)

ヒストグラム (データ分布)

相関係数のヒートマップ

ベイズ最適化

自然言語プロット

線形回帰

非線形回帰

次元削減

マニュアル

書類作成

ログアウト

キーワード検索 詳細検索

MeltingPoint 検索キーワード一覧をカンマ区切りで入力してください

②

③ 検索

- ① 「検索」をクリック
- ② にキーワードを入力
(入力しようとすると検索候補が自動表示)
⇒例えば melting point を選択
- ③ 「検索」をクリック

データ連携と検索

MIX Platform

① ファイルブラウザ

検索

ファイルサイズ

電気化学測定

プロット・解析

データの閲覧

MI解析

データ閲覧

データ修正

プロット (散布図)

ヒストグラム (データ分布)

相関係数のヒートマップ

ベイズ最適化

自然言語プロット

線形回帰

非線形回帰

次元削減

マニュアル

書類作成

ログアウト

キーワード検索 詳細検索

MeltingPoint 検索キーワード一覧をカンマ区切りで入力してください

②

③ 検索

検索結果

検索キーワード: MeltingPoint

選択したファイルをダウンロード

マージビューアを開く

☒	Filename	File Full Path	最終更新日時
☒	solvent_carbonate_mpbp.csv	/workshop2410_search_demo_project/講習会用検索デモ__1727250072159/raw_data/...	2024/09/25 07:46:36
☒	solvent_web1_mpbp.csv	/workshop2410_search_demo_project/講習会用検索デモ__1727250072159/raw_data/...	2024/09/25 07:46:36
☒	solvent_ether_mpbp.csv	/workshop2410_search_demo_project/講習会用検索デモ__1727250072159/raw_data/...	2024/09/25 07:46:36
☒	solvent_lactone_mpbp.csv	/workshop2410_search_demo_project/講習会用検索デモ__1727250072159/raw_data/...	2024/09/25 07:46:36
☒	solvent_nitrile_mpbp.csv	/workshop2410_search_demo_project/講習会用検索デモ__1727250072159/raw_data/...	2024/09/25 07:46:36

5 rows selected

Rows per page: 25 1-5 of 5

- ① 「検索」をクリック
- ② にキーワードを入力
(入力しようとすると検索候補が自動表示)
⇒例えば melting point を選択
- ③ 「検索」をクリック
- ④ melting point の項目を含むファイルが表示される。ひとまず全ファイルを選択☑

name	Molecular weight	density	Boiling Point	Melting Point	Flash point	LUMO	HOMO	Dipole moment
PC	102.09	1.21	242	-49	132	0.946	-7.930	5.255
EC	88.06	1.32	244	38	152	0.919	-8.017	5.07
VC	86.05	1.255	162	22	72	-0.137	-6.973	4.365
FEC	106.05	1.497	210	17.3	122	0.493	-8.468	4.487
DMC	90.08	1.07	90	4	17	1.115	-7.774	0.342
DEC	118.13	0.975	126	-43	25	1.217	-7.654	0.613
EMC	104.1	1.015	107	-55	23	1.168	-7.713	0.514

①「検索」をクリック
 ②にキーワードを入力
 (入力しようとすると検索候補が自動表示)

⇒例えば melting point を選択
 ③「検索」をクリック

④melting point の項目を含むファイルが表示される。ひとまず全ファイルを選択☑

- データ閲覧
- データ修正
- プロット (散布図)
- ヒストグラム (データ分布)
- 相関係数のヒートマップ
- ベイズ最適化
- 自然言語プロット
- 線形回帰
- 非線形回帰
- 次元削減
- マニュアル
- 書類作成
- ログアウト

☒	Filename	File Full Path	最終更新日時
☒	solvent_carbonate_mpbp.csv	/workshop2410_search_demo_project/講習会用検索デモ__1727250072159/raw_data/...	2024/09/25 07:46:36
☒	solvent_web1_mpbp.csv	/workshop2410_search_demo_project/講習会用検索デモ__1727250072159/raw_data/...	2024/09/25 07:46:36
☒	solvent_ether_mpbp.csv	/workshop2410_search_demo_project/講習会用検索デモ__1727250072159/raw_data/...	2024/09/25 07:46:36
☒	solvent_lactone_mpbp.csv		
☒	solvent_nitrile_mpbp.csv		

name	density	Boiling Point	Melting Point	Flash point
酢酸	0.78	118	15	40
無水酢酸	1.08	140	-73	49
アセトン	0.79	56	-94	-18
アニリン	1.021	184	-6	70

複数の表を合体させて一つの大きな表を作り、機械学習を行いたい。

データ連携と検索

MIX Platform

① ファイルブラウザ

検索

② キーワード検索

MeltingPoint × 検索キーワード一覧をカンマ区切りで入力してください

③ 検索

④ 検索結果

検索キーワード: MeltingPoint

⑤ 選択したファイルをダウンロード

マージビューアを開く

☒	Filename	File Full Path	最終更新日時
☒	solvent_carbonate_mpbp.csv	/workshop2410_search_demo_project/講習会用検索デモ__1727250072159/raw_data/...	2024/09/25 07:46:36
☒	solvent_web1_mpbp.csv	/workshop2410_search_demo_project/講習会用検索デモ__1727250072159/raw_data/...	2024/09/25 07:46:36
☒	solvent_ether_mpbp.csv	/workshop2410_search_demo_project/講習会用検索デモ__1727250072159/raw_data/...	2024/09/25 07:46:36
☒	solvent_lactone_mpbp.csv	/workshop2410_search_demo_project/講習会用検索デモ__1727250072159/raw_data/...	2024/09/25 07:46:36
☒	solvent_nitrile_mpbp.csv	/workshop2410_search_demo_project/講習会用検索デモ__1727250072159/raw_data/...	2024/09/25 07:46:36

5 rows selected

Rows per page: 25 1-5 of 5

- ① 「検索」をクリック
- ② にキーワードを入力
(入力しようとすると検索候補が自動表示)
⇒例えば melting point を選択
- ③ 「検索」をクリック
- ④ melting point の項目を含むファイルが表示される。ひとまず全ファイルを選択
- ⑤ 「マージビューアを開く」をクリック

データ連携と検索

MIX Platform

マージするデータを選択

データ一覧を表示する

マージデータのプレビュー

前の画面に戻る **⑥** エクスポート

Line	Data ID	name	Molecularweight	density	BoilingPoint	MeltingPoint	Flashpoint	LUMO	HOMO	Dipole
1	0	PC	102.09	1.21	242	-49	132	0.946	-7.93	
2	0	EC	88.06	1.32	244	38	152	0.919	-8.017	
3	0	VC	86.05	1.255	162	22	72	-0.137	-6.973	
4	0	FEC	106.05	1.497	210	17.3	122	0.493	-8.468	
5	0	DMC	90.08	1.07	90	4	17	1.115	-7.774	
6	0	DEC	118.13	0.975	126	-43	25	1.217	-7.654	
7	0	EMC	104.1	1.015	107	-55	23	1.168	-7.713	
8	1	酢酸		0.78	118	15	40			
9	1	無水酢酸		1.08	140	-73	49			
10	1	アセトン		0.79	56	-94	-18			
11	1	アニリン		1.021	184	-6	70			
12	2	DMM	76.09	0.86	42	-105	-32	1.459	-6.846	
13	2	DME	90.12	0.867	85	-58	0	1.669	-6.863	
14	2	DEE	118.17	0.842	121	-74	35	1.665	-6.839	
15	2	Diglyme	134.17	0.945	162	-64	56	1.616	-6.898	
16	2	Triglyme	178.18	0.9862	216	-45	113	1.61	-6.907	
17	3	d-Valero lactone		1.079	245	-13	112			
18	3	e-Capro lactone		1.03	253	-1	109			
19	3	g-Hexano lactone		1.03	220	-18	98			
20	4	AN		0.783	82	-45	2	0.898	-8.933	

⑥ 「エクスポート」をクリック
⇒データ連携されたcsvファイルが、ローカルPCにダウンロードされる。

←複数のファイルが、共通の項目名を持つものに対して、列を揃えて出力される。

検索

name	Molecular weight	density	Boiling Point	Melting Point	Flash point	LUMO	HOMO	Dipole moment
PC	102.09	1.21	242	-49	132	0.946	-7.93	5.255
EC	88.06	1.32	244	38	152	0.919	-8.017	5.07
VC	86.05	1.255	162	22	72	-0.137	-6.973	4.365
FEC	106.05	1.497	210	17.3	122	0.493	-8.468	4.487
DMC	90.08	1.07	90	4	17	1.115	-7.774	0.342
DEC	118.13	0.975	126	-43	25	1.217	-7.654	0.613
EMC	104.1	1.015	107	-55	23	1.168	-7.713	0.514
酢酸		0.78	118	15	40			
無水酢酸		1.08	140	-73	49			
アセトン		0.79	56	-94	-18			
アニリン		1.021	184	-6	70			
DMM	76.09	0.86	42	-105	-32	1.459	-6.846	
DME	90.12	0.867	85	-58	0	1.669	-6.863	
DEE	118.17	0.842	121	-74	35	1.665	-6.839	
Diglyme	134.17	0.945	162	-64	56	1.616	-6.898	
Triglyme	178.18	0.9862	216	-45	113	1.61	-6.907	
d-Valero lactone		1.079	245	-13	112			
e-Capro lactone		1.03	253	-1	109			
g-Hexano lactone		1.03	220	-18	98			
AN		0.783	82	-45	2	0.898	-8.933	
PN		0.772	97	-93	6	0.587	-8.802

⑥ 「エクスポート」をクリック
⇒データ連携されたcsvファイルが、ローカルPCにダウンロードされる。

←複数のファイルが結合した一つのcsvファイルが作られる

⇒これをMIXのMI解析にかける。

70			
-32	1.459	-6.846	
0	1.669	-6.863	
35	1.665	-6.839	
56	1.616	-6.898	
113	1.61	-6.907	
112			
109			
98			

20	4	AN	0.783	82	-45	2	0.898	-8.933	
----	---	----	-------	----	-----	---	-------	--------	--

データ連携と検索

⑦

The screenshot shows the MIX Platform interface. On the left is a navigation menu with various options. The 'ファイルブラウザ' (File Browser) option is highlighted with a red box, and a circled '7' is placed above it. The main area shows the 'Current Project' as 'workshop2410_private_project_01' and the 'Current Directory' as '/'. There are buttons for 'LOAD', 'DELETE', and 'DOWNLOAD' at the top right of the main area. Below these buttons is a large empty space with a small icon and text '新規フォルダを作成...' (Create new folder...).

⑦ 「ファイルブラウザ」をクリック。MI解析をするために、先ほど合体させたExcelファイル(csvファイルとしてローカルPCにダウンロードされている)をMIXにアップロードする。

データ連携と検索

The screenshot displays the MIX Platform interface. At the top, a green header bar contains the text "MIX Platform" and a notification "アップロードが完了しました。ファイルID : 35" with a close button. Below the header, a dropdown menu shows "Current Project : workshop2410_private_project_01". The current directory is "/merged_data_1727685607__1727686755927/raw_data/". A file browser shows a folder "merged_data_1727685607__1727686755927" and a file "merged_data_1727685607.csv". The file details for "merged_data_1727685607.csv" are: Type: csv, Size: 1.38 KB, Last modified: 24/09/30 17:59. Action buttons "LOAD CSV", "DELETE", and "DOWNLOAD" are visible. A sidebar on the left lists various analysis tools like "電気化学測定", "プロット・解析", "データの閲覧", "MI解析", etc.

- ⑦ 「ファイルブラウザ」をクリック。MI解析をするために、先ほど合体させたExcelファイル(csvファイルとしてローカルPCにダウンロードされている)をMIXにアップロードする。
- ⑧ Current Projectを必ずprivateの方にする。(workshop2410_private_project_*)
- ⑨ アップロードしたファイルを選択。
- ⑩ LOAD CSVをクリック

データ連携と検索

<

MIX Platform

⑪データ閲覧をクリック

ファイルブラウザ

検索

ファイルサイズ

電気化学測定

プロット・解析

データの閲覧

MI解析

データ閲覧

⑪

データ修正

プロット (散布図)

ヒストグラム (データ分布)

相関係数のヒートマップ

ベイズ最適化

自然言語プロット

線形回帰

非線形回帰

次元削減

マニュアル

書類作成

ログアウト

Data Viewer

Enabled	Name	Molecularweight	density	BoilingPoint	MeltingPoint	Flashpoint	LUMO	HOMO	Dipolemoment
✓	PC	102.09	1.21	242	-49	132	0.946	-7.93	5.255
✓	EC	88.06	1.32	244	38	152	0.919	-8.017	5.07
✓	VC	86.05	1.255	162	22	72	-0.137	-6.973	4.365
✓	FEC	106.05	1.497	210	17.3	122	0.493	-8.468	4.487
✓	DMC	90.08	1.07	90	4	17	1.115	-7.774	0.342
✓	DEC	118.13	0.975	126	-43	25	1.217	-7.654	0.613
✓	EMC	104.1	1.015	107	-55	23	1.168	-7.713	0.514
✓	酢酸		0.78	118	15	40			
✓	無水酢酸		1.08	140	-73	49			
✓	アセトン		0.79	56	-94	-18			
✓	アニリン		1.021	184	-6	70			
✓	DMM	76.09	0.86	42	-105	-32	1.459	-6.846	
✓	DME	90.12	0.867	85	-58	0	1.669	-6.863	
✓	DEE	118.17	0.842	121	-74	35	1.665	-6.839	
✓	Diglyme	134.17	0.945	162	-64	56	1.616	-6.898	

データ連携と検索

MIX Platform

Data Viewer

Enabled	Name	Molecularweight	density	BoilingPoint	MeltingPoint	Flashpoint	LUMO	HOMO	Dipolemoment
☒	PC	102.09	1.21	242	-49	132	0.946	-7.93	5.255
☒	EC	88.06	1.32	244	38	152	0.919	-8.017	5.07
☒	VC	86.05	1.255	162	22	72	-0.137	-6.973	4.365
☒	FEC	106.05	1.497	210	17.3	122	0.493	-8.468	4.487
☒	DMC	90.08	1.07	90	4	17	1.115	-7.774	0.342
☒	DEC	118.13	0.975	126	-43	25	1.217	-7.654	0.613
☒	EMC	104.1	1.015	107	-55	23	1.168	-7.713	0.514
☐	酢酸		0.78	118	15	40			
☐	無水酢酸		1.08	140	-73	49			
☐	アセトン		0.79	56	-94	-18			
☐	アニリン		1.021	184	-6	70			
☒	DMM	76.09	0.86	42	-105	-32	1.459	-6.846	
☒	DME	90.12	0.867	85	-58	0	1.669	-6.863	
☒	DEE	118.17	0.842	121	-74	35	1.665	-6.839	
☒	Diglyme	134.17	0.945	162	-64	56	1.616	-6.898	

⑪データ閲覧をクリック

⑫データ欠損が無いように、空白部分を消す。

データ連携と検索

MIX Platform

EXPORT PPTX

Scatter Plot

⑬ MI解析

⑬ データ閲覧

⑭ density

⑭ MeltingPoint

The image shows a software interface for data analysis. On the left is a sidebar menu with various options. The 'MI解析' (MI Analysis) option is highlighted with a red circle and the number 13. Below it, 'データ閲覧' (Data Browse) is also highlighted with a red circle and the number 13. The main area displays a scatter plot titled 'Scatter Plot' with 'MeltingPoint' on the y-axis and 'density' on the x-axis. Below the plot, there are two dropdown menus. The first dropdown is labeled 'density' and has a red circle with the number 14 next to it. The second dropdown is labeled 'MeltingPoint' and also has a red circle with the number 14 next to it. The plot shows several data points scattered across the axes.

⑬ MI解析のデータ閲覧をクリック

⑭ 例えば、
y軸をdensityに、
x軸をMeltingPointにする。
⇒ MeltingPointを
目的変数として予測
を実行する。

データ連携と検索

MIX Platform

⑮ 非線形回帰の Random Forest をクリックして実行。

⑮ 比較的よく Melting Point を予測できた。

⑰ Melting Point の予測には density が効く。(実験的に知られた事実)

The screenshot displays the MIX Platform interface. On the left is a navigation menu with various analysis tools. The main area shows two plots related to Melting Point analysis. Plot 17 is a bar chart titled 'feature importance of MeltingPoint' showing the weight of different features. Plot 16 is a scatter plot titled 'True value vs Prediction (MeltingPoint)' showing the relationship between true and predicted values. Below the plots is a 'Max Depth' slider and an 'APPLY' button.

Feature	Weight
density	0.45
BoilingPoint	0.08
Flashpoint	0.12
LUMO	0.25
HOMO	0.08

True value	Prediction
-45	-45
-40	-40
-35	-35
-30	-30
-25	-25
-20	-20
-15	-15
-10	-10
-5	-5
0	0
5	5
10	10
15	15
20	20
25	25
30	30
35	35
40	40
45	45

データ連携と検索

複数ファイルを連携して利用

Line	Data ID	name	Molecularweight	density	BoilingPoint	MeltingPoint	Flashpoint	LUMO	HOMO	Dist
1	0	PC	102.0	1.21	242	-49	132	0.946	-7.93	
2	0	EC	88.0	1.32	244	38	152	0.919	-8.017	
3	0	VC	86.0	1.255	162	22	72	-0.137	-6.973	
4	0	FEC	106.0	1.497	210	17.3	122	0.493	-8.468	
5	0	DMC	90.0	1.07	90	4	17	1.115	-7.774	
6	0	DEC	118.1	0.975	126	-43	25	0.217	-7.654	
7	0	EMC	104.1	1.015	107	-55	23	1.161	-7.713	
8	1	酢酸		0.78	118	15	40			
9	1	無水酢酸		1.08	140	-73	49			
10	1	アセトン		0.79	56	-94	-18			
11	1	アニリン		1.021	184	-6	70			
12	2	DMM	76.0	0.86	42	-105	-32	1.459	-6.84	
13	2	DME	90.12	0.867	85	-58	0	1.669	-6.863	
14	2	DEE	118.1	0.942	121	-74	35	1.665	-6.839	
15	2	Diglyme	134.1	0.945	162	-64	56	1.616	-6.898	
16	2	Triglyme	178.1	0.962	216	-45	113	1.61	-6.907	
17	3	d-Valero lactone		1.079	245	-13	112			
18	3	e-Capro lactone		1.03	253	-1	109			
19	3	g-Hexano lactone		1.03	220	-18	98			
20	4	AN		0.783	82	-45	2	0.898	-8.933	
21	4	PN		0.772	97	-93	6	0.587	-8.802	
22	4	Adiponitrile		0.96	295	1	93	0.312	-8.976	
23	4	Valeronitrile		0.8	140	-96	40	0.728	-8.704	
24	4	Glutaronitrile		0.989	286	-26	109	-0.136	-9.081	
25	4	Malononitrile		1.049	220	32	112	-0.51	-9.544	
26	4	Succinonitrile		1.02	267	56	158	-0.097	-9.31	
27	4	Suberonitrile		0.93	172	-3.5	161	0.497	-8.789	
28	4	Isobutyronitrile		0.77	107.5	-72	8	0.642	-8.727	

MIX Platform

feature importance of MeltingPoint

True value vs Prediction (MeltingPoint)

多数のデータで精度よく予測
densityという特徴量を正しく選択

1ファイルのみのデータを利用

MIX Platform

feature importance of MeltingPoint

True value vs Prediction (MeltingPoint)

Max Depth

少数データでの予測となり、信頼度が低い
特徴量の選択も信頼性が下がる

まとめ

1. PEFCに特化したデータ解析・蓄積プラットフォーム（MIX Platform）を構築した。
2. 本日は、電気化学測定データの解析と、MI解析部分に絞って説明した。
3. 今後、機能のアップデートを順次実施。
4. 企業、アカデミアの方のご利用をお待ちしています。
 - ・アカデミアの方、ご連絡頂ければすぐにご利用可能です。
 - ・企業の方、L10後すぐにご利用可能です。

MI解析の勘所 (MIノウハウ)

- どのような問題に対して、どのような手法を使えばよいか？

解析の目的	データ数	MI手法	例
1. 要因解析	少 多	回帰分析 Deep Learning	膜劣化の要因解析
2. (人に認識できない) 要因解析	中(スペクトル)	次元削減+回帰分析	膜劣化の要因解析
3. (実験が大変な) 物性予測	少 多	回帰分析 Deep Learning	触媒活性、耐久性予測
4. 実験条件最適化	少 多	ベイズ最適化 Deep Learning	触媒作成条件の最適化
5. 新物質提案 (組成) 新物質提案 (分子式)	少 中	ベイズ最適化 Deep Learning (SMILES-X)	新規触媒組成の提案 高分子膜材の提案