2017/7/10

# MI²I 第１回　Python を使ったマテリアルズ・インフォマティクス講習会報告

主催　NIMS 情報統合型物質・材料開発イニシアティブ
科学技術振興機構
計算科学振興財団
計算物質科学人材育成コンソーシアム

開催日：2017 年 7 月 7 日（金）13:00-16:00 （12:30 受付開始）
会場：高度計算科学研究支援センター　実習室
（〒650-0047 神戸市中央区港島南町 7 丁目 1 番 28 号　計算科学センタービル）

材料開発・研究に情報科学を取り入れるマテリアルズ・インフォマティクスが注目されている。本講習会は NIMS 情報統合型物質・材料開発イニチアチブ（MI²I）を中心にした四機関による共同開催であり、端末にインストールされた Python ベースのデータ解析ソフトとテスト用データセットを使った演習から、データ駆動の基礎を理解することを目指している。 7 月 7 日と 9 月 29 日の二回シリーズと 1 月 19 日の続編を企画しており、ここでは 7 月 7 日に実施された第 1 回の講習会について報告する。

第一回（7/7）プログラム

| | | |
|---|---|---|
| 12:30 | 受付開始 | |
| 13:00- | ・ Python 説明（座学）<br>・回帰分析（端末を使った演習）<br>・次元削減（端末を使った演習）<br>（適宜休憩を挟みながら） | 講師：木野日織<br>＋アシスタント |
| 16:15 | 終了 | |

1. 講師：NIMS　木野日織
2. コーディネータ・進行：NIMS　石井真史
3. TA：大阪大学　黒田文彬　福島鉄也
   NIMS　山下智樹　藤井将　青木祐太（敬称略）
4. 参加者：20 名（定員 20 名）　NIMS：真鍋 明、篠田陽子、JST：木原尚子
   （今回講習会は実習であることから、参加条件として、各自 PC に Anaconda4.4.0 事前インストールし持参、jupyter notebook が操作できること、サンプルコード＆サンプルデータ：bitibucket より事前にダウンロードとした。）

５．総括コメント

・事前に anaconda インストールと資料ダウンロードを参加条件としたこともあり、プログラム実行上のトラブルはほぼ皆無。

・ipynb 形式の実習資料は、プレゼン資料であり、学習テキストであり、同時にサンプルコードでもあるためハンズオンに有望な手段といえる。それと同時にプログラムを書き下すことなく実行できてしまうため、手を動かす実習となりにくい側面がある。

・希土類遷移金属のキュリー温度データセットを用いた解析事例コードなど講師の事前準備により、マテリアルインフォマティクスをリアルに実習する機会となった。

６．主な学習内容：

事前ダウンロードのサンプルコード ipynb を用いプログラムの概要を説明しつつ進める形式とした。

・データ解析の手順解説

・導入編として人工データによる線形回帰学習　LASSO、ridge

・実データ（キュリー温度に関する論文データ）よる回帰実習

・主成分分析（PCA）法の事例学習

・全状態探索（Exhaustive search）による重要記述子の抽出事例学習

７．主な質疑応答

学習テキストを超えたデータ解析に関する質問が多くあった。

例）多目的最適化（パレート最適)、巨大記述子群の次元削減（SIS+Lasso)、ハイパーパラメータ最適化、Cross-validation(k-fold 法と leave-one-out 法）など

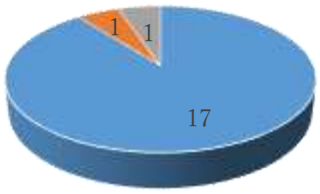
８．講習会終了後、FOCUS 小八木様のご厚意で、約１時間　J-FOCUS 施設と理研「京」見学。「京」本体見学は予約が必要との事で今回は「京」建屋１F のみ。

９．会場の様子（写真）

10.　アンケートから（回収数 19）

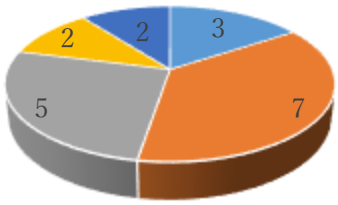
■参加者バックグラウンド

参加者所属種別



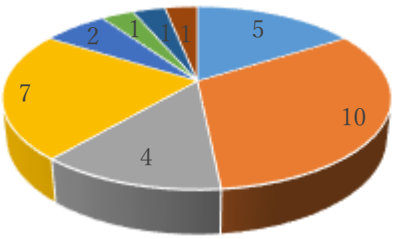

1.民間企業
2.公設機関（研究所、官公庁など）
3.教育機関（大学等）

年齢



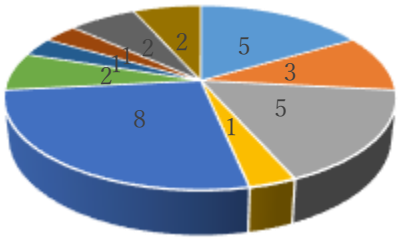

20代　30代　40代　50代　60代

専門分野（複数回答可）




1.物理　2.化学　3.無機材料
4.有機材料　5.機械　7.生物
9.電気・電子　14.その他

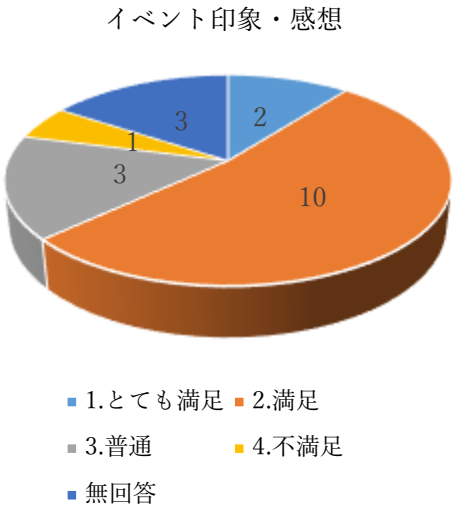
関係する産業分野（複数回答可）

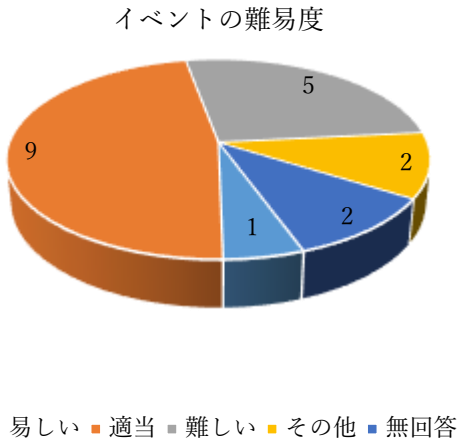



2.自動車　4.輸送用機器
5.繊維　6.精密機器
7.化学　13.ゴム
17.鉄鋼　19.機械
20.電気機器　25.非鉄金属製品

今回参加者は殆どが民間企業の所属で、30 代、40 代の中堅層。専門または関連する産業分野は化学系（有機・繊維含む）が多いが、その他輸送機器（自動車）、鉄鋼など MI に関心を寄せる産業分野の幅が伺える。

■イベント印象・感想など




アンケート回答者の半数以上で「満足」以上の評価を頂いた。

難易度に関しては回答者の半数から適当以上の回答を頂いたが、参加者の 1/4 から「難しい」との回答が寄せられている。「適当」「難しい」のいずれの回答者からも、事前の送付資料により予習ができる点を評価するコメントが寄せられ、出来る限り継続が望まれる。

■今後行ってほしいイベント（寄せられたご意見よりチュートリアルに絞って記載）

Python 演習

・Python の実習（応用を含む）
・Python の基礎的講習会。理論の演習。
・Python×機械学習で基礎固めの講習会
・実際的な演習とその解説

演習対象

・画像認識
・"ニューラルネットワーク、（ディープラーニング）を用いた実践形式講習会
・多次元尺度構成法（MDS）等、実際に材料開発プロセスで活用できる手法についての実践形式講習会
・非線形回帰。データの前処理。逆構造解析。欠損値の扱い（ノイズ、外れ値処理）
・高分子分野を例題にした解析についての講習会
・Real な系を使った事例での実習

MI への応用

・マテリアルズ・インフォマティクスにつながる Python による機械学習
・DPF 活用のための webAPI 作成の講習会
・DPF 使用のための初歩的な Linux の講習会
・具体的な機械学習の使い方を体感したい。（便利な手法、実際に使われている手法）
・MI の戦法の立て方（事例を交えた、具体的な進め方の MI の技術講習）

その他

・第一原理計算の実習。
・Advanced コースとして、論文事例の追体験。
・今回のような形式のハンズオンのシリーズ化。

以上