

2017/9/29

MI²I 第2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9月29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・材料開発イニシアティブ（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9月29日に実施された第2回の講習会について報告する。

第二回（9/29）プログラム

12:30	受付開始	
13:00-	・Python 復習（座学） ・分類（端末を使った演習） ・クラスタ分析（端末を使った演習） ・総括 (適宜休憩を挟みながら)	講師：木野日織 +アシスタント
16:15	終了	

(1)参加者（敬称略）

1. 講師：NIMS 木野日織
2. 進行：JST 木原尚子
3. TA：大阪大学 黒田文彬、神田洋佑
NIMS 青木祐太
4. スタッフ：NIMS 篠田陽子
5. 受講者：19名（定員20名）

（前回に続き、参加条件として、各自PCにAnaconda4.4.0を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。)

(2)実施内容・プログラム：

機械学習の基本的な手法（次元削減、クラス分類、クラスタリング）に関して物性の基本データを用いて記述子作成のアルゴリズムおよび解析方法について演習した。

・演習 1：クラス分類に関しては、Luca M.等の論文を元に様々な物性値から、2 元系材料の結晶構造を structures の energy 差を線形回帰して結晶構造に分類する演習問題。データの規格化、記述子の抽出を行い、訓練データとテストデータ（90：10）に分け、L1 正則化、L2 正則化の 2 種類のモデルによる解析、ハイパーパラメーターの調整法、モデルの評価法などを学習した。

・演習 2：クラス分類の 2 つ目の演習として、元素の物性値から結晶構造 (b.c.c, f.c.c, h.p.c) を予測する演習。基本データ（酸化価数の最小と最大、原子半径、電子陰性度など）としては wikipedia からデータが揃うものを用いている。これらから、記述子を抽出し、L1、L2 正則化モデルを使って結晶構造予測を行った。

・演習 3：クラスタリングの演習。カーボンの 3 次元構造体の類似性を分類する手法としてクラスタリングモデルを適用し演習課題とした。

(写真)

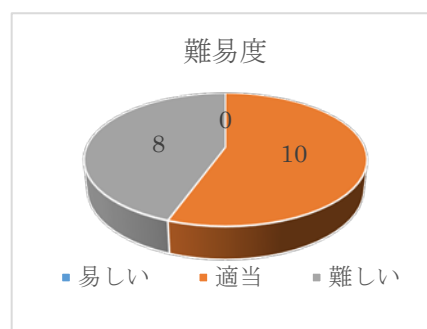


(3)アンケートから

前回同様、受講者アンケートを実施した。

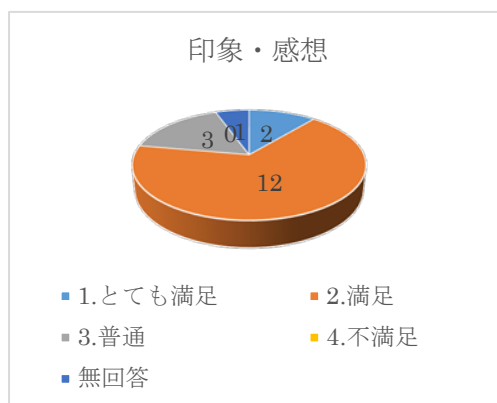
参加者 19 名中 18 名より回答を頂いた。

講習会の難易度に関しては、第 1 回に比較し難しいとの回答（前回 5 名）が増えていたが、印象・感想では前回（12 名）を上回る 14 名から満足以上の回答



が寄せられており講習会全体は高い評価を頂いている。

また、毎回木野講師に作成いただいている教材は受講者には事前学習教材としても好評であり、今後、教材を活用したセミナーを企画していきたい。1月19日の続編企画への積極的な要望記載も多く、頂いた意見を活用していく。



以上