

平成30年度

社会人向け実践教育プログラム
「先端データサイエンス実践講座」

日程およびシラバス

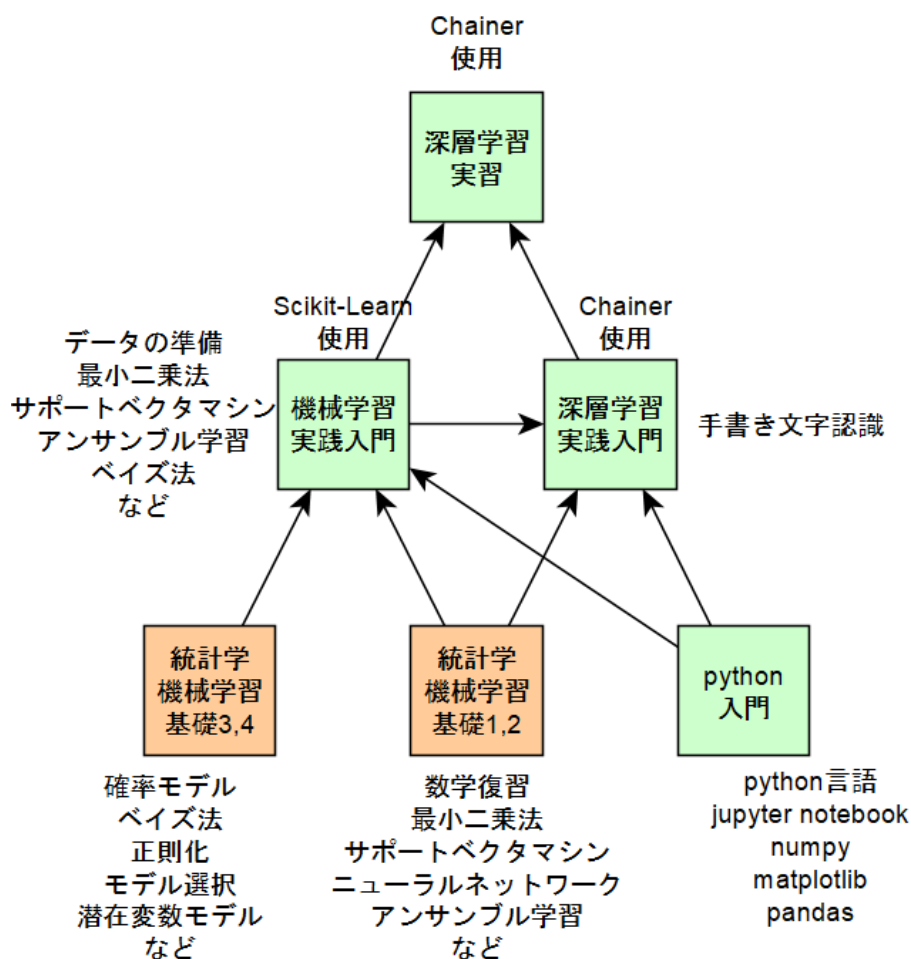
豊橋技術科学大学
社会連携推進センター

目次

1	はじめに	3
2	日程 (全 45H).....	4
3	シラバス	5
3.1	統計学・機械学習基礎 1,2 (6H)	5
3.2	統計学・機械学習基礎 3,4 (6H)	5
3.3	python 入門 1,2 (9H)(オプション)	5
3.4	機械学習実践入門 (4.5H)	6
3.5	深層学習環境準備・入門 (4.5H)	6
3.6	深層学習実習 (9H)	6
3.7	先端データサイエンス講演会 (6H)	7
4	その他情報.....	7
4.1	e-Learning について	7
5	計算技術科学特論 (全 15H, 30 講義から 10 講義を選択)	8
5.1	計算科学特論 A (2017).....	8
5.2	計算科学特論 B (2018).....	9

1 はじめに

今日、コンピュータおよびその周辺技術の急速な進展により、膨大なデータを取得、生成することが可能な時代になりました。このビッグデータの中に埋もれた有用な情報を取り出し、活用するためにはデータサイエンスの基礎理論から実践的ノウハウまで広く理解しておく必要があります。豊橋技術科学大学は、開学以来、「技術」を「科学」で裏付けし、そこから新しい技術を創造する技術科学の教育・研究を使命としています。とくに、本学の知識・情報工学系において、これまで「情報化学」などの様々なデータサイエンスの応用分野で教育・研究を行ってきた実績があります。これらのノウハウを本学学生のみならず、これからデータサイエンスの応用を目指そうとする社会人技術者にも提供したいと考えています。本コースでは、統計学や機械学習などのデータサイエンスの基礎理論から、「計算物質科学」等への機械学習の実践応用まで学ぶことができます。



2 日程 (全 45H)

回	タイトル	担当	日にち	時限	概要
1	統計学・機械学習基礎 1	金澤	10/3(水)	3,4	機械学習・パターン認識論
2	統計学・機械学習基礎 2	金澤	10/3(水)	5,6	機械学習・パターン認識論
3	統計学・機械学習基礎 3	渡辺	10/10(水)	3,4	機械学習の数学的基礎
4	統計学・機械学習基礎 4	渡辺	10/10(水)	5,6	機械学習の数学的基礎
5	Python 入門 1 (オプション)	濱田	10/17(水)	3,4,5	python 言語, jupyter notebook, numpy, matplotlib
6	Python 入門 2 (オプション)	濱田	10/24(水)	3,4,5	pandas, 簡単な python プログラミング
7	機械学習実践入門	濱田	11/7(水)	3,4,5	機械学習のプログラミング 例題(Scikit-Learn を使用)
8	深層学習環境準備	濱田	11/7(水)	6	深層学習フレームワークである Chainer の環境準備
9	深層学習実践入門	濱田	11/21(水)	3,4	Chainer+Cupy の基本概念、 手書き文字認識の例題
10	深層学習実習 1	濱田	11/21(水)	5,6	Scikit-Learn の例題の Chainer 版への置き換え
11	深層学習実習 2	濱田	11/28(水)	3,4	再帰型ニューラルネットワークと簡単な自然言語処理
12	深層学習実習 3	濱田	11/28(水)	5,6	分子活性予測
13	先端データサイエンス講演会	外部講師	未定	2,3,4,5	マテリアルズインフォマティクス等

[コメント]

時限定義

2: 10:30-12:00, 3:13:00-14:30, 4:14:40-16:10, 5:16:20-17:50, 6: 18:00-19:30 を意味する。

☐

3 シラバス

3.1 統計学・機械学習基礎 1.2 (6H)

タイトル	担当	日にち	備考
統計学・機械学習基礎 1,2	金澤靖	10/3	
内容			
機械学習とパターン認識に必要となる統計学や線形代数の基礎知識を学び、基本的な手法であるサポートベクトルマシン、ニューラルネットワーク、アンサンブル学習などの基本原理を学ぶ。。			

3.2 統計学・機械学習基礎 3,4 (6H)

タイトル	担当	日にち	備考
統計学・機械学習基礎 3,4	渡辺一帆	10/10	
内容			
機械学習で重要となる潜在変数を持つ学習モデルやベイズ推測の枠組みを学ぶ。高専・大学初等の数学的知識を用いた、学習法の導出を通じて、確率モデルを用いた機械学習への理解を深める。			

3.3 python 入門 1,2 (9H)(オプション)

タイトル	担当	日にち	備考
python 入門 1,2	濱田信次	10/17, 10/24	
内容			
データサイエンスで中心的に使用される言語である python 言語(ここでは python3 を使用)の基本的使い方を習得する。(資料の多くを jupyter notebook 形式で提供予定のため、その簡単な説明をあらかじめ行う。) また python におけるもっとも多用されるパッケージである numpy(配列処理)と matplotlib(可視化)の基本的使い方を習得する。さらにはデータサイエンスにおいて多用されるパッケージである pandas も習熟し、簡単なプログラミング手法を理解する。python 入門を受講するかどうかは選択可能としますが、python 言語の基礎を習得されていない場合、以後の講義・実習に支障をきたしますので、かならず受講をしてください。			

3.4 機械学習実践入門 (4.5H)

タイトル	担当	日にち	備考
機械学習実践入門	濱田信次	11/7	
内容			
Scikit-Learn は使い勝手の良い、機械学習用 Python パッケージである。Scikit-Learn を利用したいいくつかのサンプルプログラムを動かすことによって、「統計学・機械学習基礎」で学んだ内容をより具体的に理解すると同時に Chainer 入門へつなげる。			

3.5 深層学習環境準備・実践入門 (4.5H)

タイトル	担当	日にち	備考
深層学習環境準備・実践入門	濱田信次	11/7, 11/21	
内容			
最初に Chainer 利用環境(Linux)の準備を行う。その後、Chainer のサンプルプログラムである手書き文字認識を例題として、Chainer の基本概念、利用方法について理解する。それと同時にニューラルネットワークの動作原理についても理解する。 さらに、chainer から利用される cupy (Python から cuda を簡単に利用するためのライブラリ) の基本概念、利用方法についても理解する。			

3.6 深層学習実習 (9H)

タイトル	担当	日にち	備考
深層学習実習 1,2,3	濱田信次	11/21, 11/28	
内容			
深層学習実習 1 では Scikit-Learn 入門で使用したサンプルプログラムを Chainer 版に置き換えることにより、Chainer の利用方法の理解を深める。深層学習実習 2 では再帰ニューラルネットワーク(フィードバックループをもつニューラルネットワーク) の簡単な例題の説明をおこなったのち、自然言語処理への応用例について学習する。深層学習実習 3 では分子活性予測の簡単な例を取り上げる。分子活性予測は、ディスクリプタと呼ばれる実験的あるいは理論的に得られた多くの数値データから、分子のもつ様々な活性値を予測するものであり創薬における非常に重要な技術である。取り上げる例は限られているが、データの準備方法、解析評価手法、予測性能を高めるためのパラメータ最適化などできるだけ一般的な場合にも通用する技術の習得を試みる。			

3.7 先端データサイエンス講演会 (6H)

タイトル	担当	日にち	備考
先端データサイエンス講演会	濱田信次	1/23(水)(仮)	
内容			
先端データサイエンス分野で活躍されている学内外講師を招き講演をしていただく。 詳細内容は現時点で未定ですが決定次第お知らせします。(マテリアルズインフォマティクス関係を予定しています。)			

4 その他情報

4.1 e-Learning について

e-Learning とは教育用 Web システム(Moodle)を通じて学習することです。Moodle 上には講義資料、講義ビデオ、サンプルプログラム、小テストなどが置かれており、登録ユーザーは自宅等から任意の時間にアクセスして学習することができます。本コースはすべて来学していただき受講されることを原則としますが、どうしても来学できなかった受講者は後日 e-Learnig で学習してください。

5 計算技術科学特論（全 15H, 30 講義から 10 講義を選択）

本「先端データサイエンス実践コース」受講者に対しては、さらに、「計算技術科学特論」を受講していただくことも可能です。（追加の費用はかかりません。）

タイトル	担当	日にち	備考
計算技術科学特論	後藤仁志 濱田信次	随時	e-Learning
内容			
大阪大学から配信された計算科学技術特論 A,B（全 30 講義）から受講者が興味がある 10 講義(15 時間分) を視聴し、計算技術科学に関するさまざまな技術・理論を修得する。視聴した際の疑問等については、担当教官と議論して解消する。各講義について小テストも用意しておりこれに取り組むことで理解の助けとする。			

各講義の内容は以下の通りで。

5.1 計算科学特論 A (2017)

回	タイトル	担当	日にち	時限	概要
1	プログラム高速化の基礎	後藤・濱田	随時	1 時 限 分	理研 WEB 参照のこと http://www.aics.riken.jp/library/event/tokurona_170406.html
2	MPI の基礎	後藤・濱田	随時	同上-	同上
3	OpenMP の基礎	後藤・濱田	随時	同上	同上
4	Hybrid 並列化技法	後藤・濱田	随時	同上	同上
5	プログラム高速化の応用	後藤・濱田	随時	同上	同上
6	線形代数演算ライブラリ BLAS と LAPACK の基礎と実践 1	後藤・濱田	随時	同上	同上

7	線形代数演算ライブラリ BLAS と LAPACK の基礎と実践 2	後藤・濱田	随時	同上	同上
8	高速化チューニングとその関連技術 1	後藤・濱田	随時	同上	同上
9	高速化チューニングとその関連技術 2	後藤・濱田	随時	同上	同上
10	行列計算における高速アルゴリズム 1	後藤・濱田	随時	同上	同上
11	行列計算における高速アルゴリズム 2	後藤・濱田	随時	同上	同上
12	古典分子動力学法の高速度化	後藤・濱田	随時	同上	同上
13	Parallelization of Molecular Dynamics(分子動力学法の並列化)	後藤・濱田	随時	同上	同上
14	量子化学計算の大規模化 1	後藤・濱田	随時	同上	同上
15	量子化学計算の大規模化 2	後藤・濱田	随時	同上	同上

5.2 計算科学特論 B (2018)

回	タイトル	担当	日にち	時限	概要
1	スーパーコンピュータとアプリケーションの性能	後藤・濱田	随時	1 時 限 分	理研 WEB 参照のこと http://www.r-ccs.riken.jp/library/event/tokuronB_180406.html

2	アプリケーションの性能最適化 1 (高並列性能最適化)	後藤・濱田	随時	同上	同上
3	アプリケーションの性能最適化 2 (CPU 単体性能最適化)	後藤・濱田	随時	同上	同上
4	アプリケーションの性能最適化の実例 1	後藤・濱田	随時	同上	同上
5	アプリケーションの性能最適化の実例 2	後藤・濱田	随時	同上	同上
6	大規模系での高速フーリエ変換 1	後藤・濱田	随時	同上	同上
7	大規模系での高速フーリエ変換 2	後藤・濱田	随時	同上	同上
8	オーダーN 法 1	後藤・濱田	随時	同上	同上
9	オーダーN 法 2	後藤・濱田	随時	同上	同上
10	大規模 MD 並列化の技術 1	後藤・濱田	随時	同上	同上
11	大規模 MD 並列化の技術 1	後藤・濱田	随時	同上	同上
12	大規模量子化学計算 1	後藤・濱田	随時	同上	同上
13	大規模量子化学計算 2	後藤・濱田	随時	同上	同上
14	OpenACC ・ CUDA による GPU コンピ	後藤・濱田	随時	同上	同上

	ユーティング				
15	イ ン テ ル Xeon Phi コ プロセッサー 向け最適化、 並列化概要	後藤・濱田	随時	同上	同上