

先端データサイエンス実践コース

# 先端データサイエンス講演会

日時: 2018年1月24日(水)10時30分~17時50分

会場: 豊橋技術科学大学 情報メディア基盤センター 第1端末室

主催: 豊橋技術科学大学 電気・電子情報工学系/情報・知能工学系/環境・生命工学系

共催: 豊橋技術科学大学 社会連携推進センター/先端バイオリサーチセンター/エレクトロニクス先端融合研究所

<プログラム> 基調講演『大量のテキストデータから見えてくるもの』

井佐原 均 (豊橋技術科学大学 情報メディア基盤センター)

講演 I 『自然言語処理における深層学習の進展』

岡崎 直観 (東京工業大学 情報理工学院)

講演 II 『行動認識・理解のためのデータ解析』

秋月 拓磨 (豊橋技術科学大学 機械工学系)

講演 III 『第一原理計算による分子データベース構築と機械学習による物性予測』

中田 真秀 (理化学研究所 情報基盤センター)

<対象者> 一般(大学理工系学部卒業程度以上が望ましい)、本学教職員、本学学生

<定員> 70名程度(一般参加者を優先します)

<参加費> 一般: 2000円(当日受付にて現金でお支払いください)

本学教職員、本学学生: 無料

<概要>

豊橋技術科学大学では、スパコンを含む高速コンピュータを利活用できるシミュレーション技術者の育成に取り組んでいますが、同様に高速コンピュータを活用する新しい科学として、「データサイエンス」や「機械学習」に大きな注目が集まっています。それは、インターネット・センサー・シミュレーション等から生成される多種多様の大量データと対峙し、そこからヒトの経験や英知では見いだせない相関・因果関係を導き出すための科学技術であり、既に、ロボット(人工知能)やビジネス(マーケティング)などの様々な分野で実践応用が始まっています。そこで本学では、新しい時代を担う「技術科学」人材に必須となる「データサイエンス」について、発展的な応用事例や研究事例から理解を深める機会として、『先端データサイエンス講演会』を開催いたします。なお、本講演会は、「先端データサイエンス実践コース」の一環にもなっています。

【申込方法】

別紙「参加申込書」に必要事項を明記の上、E-mail(研究支援課 社会連携支援室 jinzai@office.tut.ac.jp 0532-81-5188)にてお申込みください。尚、関係資料は下記サイトからダウンロード可能です。

<http://www.sharen.tut.ac.jp/news/1003.html> (社会連携推進センターホームページ 新着情報)

申込期限: 1月19日(金)まで (学内は1月23日(火)まで)

# 開講スケジュール

日程	時間	科目名	講師
1月24日 (水)	10:30～12:00	大量のテキストデータから見えてくるもの	井佐原 均 豊橋技術科学大学 情報メディア基盤センター教授
	13:00～14:30	自然言語処理における深層学習の進展	岡崎 直観 東京工業大学 情報理工学院教授
	14:40～16:10	行動認識・理解のためのデータ解析	秋月 拓磨 豊橋技術科学大学 機械工学系助教
	16:20～17:50	第一原理計算による分子データベース構築と機械学習による物性予測	中田 真秀 理化学研究所 情報基盤センター技師

## 【講演の概要】

### ■大量のテキストデータから見えてくるもの

コンピュータに言葉を「理解」させようとする自然言語処理の研究は、初期の規則による手法から、大規模なテキストデータを用いる統計的な手法へと変化しました。近年、さらに大規模なデータとニューラルネットの組み合わせにより、機械翻訳をはじめとして多くの自然言語処理応用システムの精度が大幅に向上しました。ここでは、このような流れを概観し、そこで得られた知見と、いまだに得られない情報、失われてしまった情報について検討します。また、このような技術を実社会でのサービスに活用する場合の課題についても触れます。

### ■自然言語処理における深層学習の進展

深層学習は、画像処理や音声処理で大成功を収め、今の人工知能ブームの火付け役となりました。一方、言語処理では記号(例えば単語や文字)による特徴記述がある程度成功していたため、深層学習の破壊力は限定的でした。その後、言語処理でも研究成果の蓄積が進み、ほぼ全てのタスクで深層学習およびニューラル・ネットワークに基づく手法が最高性能を達成するようになりました。本発表では、自然言語処理にどのように深層学習が導入され、広まっていったのか、という流れを説明します。具体的には、分散表現、エンコーダー・デコーダー、アテンション・メカニズムなど、言語処理における深層学習の重要なアイデアを復習します。その後、機械翻訳、質問応答、対話文生成、要約文生成、評判分析、知識獲得、含意関係認識など、様々な言語処理タスクでの応用事例を紹介いたします。

### ■行動認識・理解のためのデータ解析

近年、加速度センサやジャイロセンサなどの装着型センサを用いた「行動センシング」の研究が活発です。これらの研究では、センサデバイスから取得した多変量のセンサデータをパターン認識、あるいは機械学習の手法を用いて解析することで、ユーザ個々の動きや状態、作業状況などを認識・理解することを目指しています。本講演では、その中で基盤となる行動認識技術を中心にその基本といくつかの応用事例をご紹介します。

### ■第一原理計算による分子データベース構築と機械学習による物性予測

近年機械学習が注目されているが、化学の研究や開発に於いても化学的直感が重視されているようにこれまで人がやってきた、多くの知識を元にした汎化がコンピュータでもできるようにになれば、さらに研究開発が進むに違いない。多くの知識は、データベース構築によって、また汎化は機械学習によってできるようになるのではと考えていて、それについて紹介する。