

授業科目名 (英訳)		情報分析・管理論 Information Analysis and Management				担当者所属 職名・氏名		情報学研究科 准教授 浅野 泰仁 情報学研究科 助教 加藤 誠			
配当学年	2・3 回生	単位数	2	開講年度 開 講 期	H27 前期	曜時限	月 4(5)	授業形態	講義	使用言語	日本語
[授業の概要・目的]											
<p>今日、文理の分野を問わず最先端の研究を進めていくには、大規模データの分析・管理技術が不可欠となっている。本講義では、コンピュータで問題を解くのに必要なモデル化やアルゴリズムとその情報検索への応用、データの管理およびデータから有用な情報や知識を抽出する技術、情報の分析結果の可視化と対話技術など、多くの分野で利用されているトピックを精選して講述する。特に、具体的な問題に対してどのようにそれらの技術が適用できるかを学ぶことで技術の基本的な考え方を理解し、各自の分野で応用が可能になることを目指している。</p>											
[到達目標]											
<p>上に述べたように、大規模データの分析・管理技術はあらゆる分野の研究に必要とされている。本授業では、それらの技術の仕組みの基礎を理解するとともに、「情報分析・管理演習」と連携して、各自が具体的な問題に対して技術を実践できるレベルに到達することを目標としている。</p>											
[授業計画と内容]											
○ガイダンス (1 回) :											
講義全体の概要を説明する。「情報分析・管理演習」では、プログラミング言語 R のインストール、簡単な使い方について学ぶ。											
○問題のモデル化と問題の解き方、情報検索(5 回)(担当: 浅野) :											
コンピュータで問題を解くために必要となる、問題のモデル化と、問題の解き方、すなわち、アルゴリズムについて学ぶ。具体的には、まず、グラフを用いたモデル化、アルゴリズムとその戦略(計算量、近似、動的計画など)、情報検索に用いられるランキング手法 PageRank や HITS などの応用について講述する。さらに、正規表現などの形式言語と、オートマトンや隠れマルコフモデルが生物情報学の配列解析に応用された事例を通して、情報科学の他分野への応用について学ぶ。「情報分析・管理演習」では、講義で扱った、最短経路や PageRank、配列解析・隠れマルコフモデル等を、R で計算する方法を学ぶ。また、可視化などを通じて結果を分析する基礎を身につける。											
○データマイニング (3 回) (担当: 浅野) :											
データを分析して、有用な情報や知識を抽出するための手法について講述する。具体的には、データマイニング技術(相関ルール、クラスタリング、決定木、サポートベクターマシンなど)と、そこで利用されている機械学習やベイズ推定の技術を学ぶ。「情報分析・管理演習」では、R で実行する方法を学ぶとともに、得られた結果を解析する。											
○データベース (2 回) (担当: 加藤) :											
大規模な情報を管理するデータベース技術について講述する。「情報分析・管理演習」では、データベースを MySQL を用いて実現する方法を学ぶ。											
○情報検索(2 回) (担当: 加藤) :											
大量の文書データから必要な情報を検索する技術について後述する。基礎的な検索モデル、索引技術やランキング学習について学ぶことで、現状の検索システムがどのように動作しているのかを理解する。また、検索システムの性能を定量的に評価する方法について理解を深めることにより、より良い検索システムを選択・構築する方法を身につける。											
○情報の可視化と対話技術(2 回) (担当: 加藤) :											
分析結果などの情報の可視化、インタラクション技術、情報発信技術を取り上げる。「情報分析・管理演習」では、Processing を用いてこれらを実現する方法を学ぶ。											

〔履修要件〕

本科目を履修する場合は、「情報分析・管理演習(情報学研究科開講科目)」(前期・月・5時限)*を履修すること。履修登録については別途周知する。なお、同演習は各自のノート PC 等の持ち込みを前提としているが、貸出用の PC も用意している。

〔成績評価の方法・観点及び達成度〕

出席状況および毎週のレポートによって[到達目標]の項で説明した目標に到達しているかどうかを検証し、成績を算出する。

〔教科書〕

特になし

〔参考書等〕

Jon Kleinberg, Eva Tardos (著), 浅野孝夫, 浅野泰仁, 小野孝男, 平田富夫 (翻訳) Jon Kleinberg and Eva Tardos 『アルゴリズム・デザイン Algorithm Design』 (共立出版 Addison Wesley.)

Richard Durbin, Sean R. Eddy, Anders Krogh, Graeme Mitchison. 『Biological sequence analysis-Probabilistic models of proteins and nucleic acids-』 (Cambridge University Press)

C.M. ビショップ(著), 元田浩, 栗田多喜夫, 樋口知之, 松本裕治, 村田昇 (翻訳) 『パターン認識と機械学習-ベイズ理論による統計的予測-上・下巻 Pattern Recognition and Machine Learning』 (シュプリンガー・ジャパン Springer-Verlag.)

〔授業外学習 (予習・復習) 等〕

必要な場合は授業中に指定する。

〔その他 (オフィスアワー等) 〕

自身の研究に、種類は問わず、何らかのデータを用いている学生を歓迎する。

オフィスアワーについては特に指定せず、メールのやり取りで随時行う。

浅野泰仁: asano@i.kyoto-u.ac.jp

加藤誠: kato@dl.kuis.kyoto-u.ac.jp