

科目ナンバリング		G-LAS12 80020 SJ13							
授業科目名 <英訳>	情報分析・管理演習 Information Analysis and Management, Exercise				担当者所属 職名・氏名	情報学研究科 教授	山本	章博	
						国際高等教育院 教授	田村	寛	
						非常勤講師	増田	央	
						情報学研究科 特定助教	東風上	奏絵	
						情報学研究科 特定助教	HUANG	Yin Jou	
群	大学院横断教育科目群		分野(分類)	統計・情報・データ科学系		使用言語	日本語		
旧群		単位数	1単位	週コマ数	1コマ	授業形態	演習(対面授業科目)		
開講年度・ 開講期	2024・後期		曜時限	月5		配当学年	大学院生	対象学生	全学向
(情報学研究科の学生は、全学共通科目として履修登録できません。所属部局で履修登録してください。)									
【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ため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発見するマイニング技術、情報の分析結果の可視化とインタラクション技術など、様々な分野で利用されているトピックを精選して、演習形式で具体的な問題を計算機で解く方法を学ぶ。</p>									
【到達目標】									
<p>【研究科横断型教育の概要・目的】</p> <p>上述したように、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・管理論」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】									
<p>・ガイダンス(1回) 演習の準備およびプログラミング言語Rのインストール、簡単な使い方</p> <p>・問題のモデル化と問題の解き方(2回) 講義で扱った、グラフを用いたモデル化とアルゴリズム(オイラー閉路、最短経路)やPageRankについて、Rで実際に問題を解く方法を学ぶ。また、問題を解くだけでなく、可視化などを通じて結果を分析する基礎を身につける。</p> <p>・データマイニング(6回) データを分析して、有用な情報や知識を発見する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、深層学習など)を、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。</p> <p>・データベース・情報検索・情報可視化(3回) 大規模な情報を管理するデータベース技術、大量の文書データから必要な情報を検索する技術、情報システムを利用者からのフィードバックに基づき改善する技術について、研究での応用事例も交えながら、Python, Jupyter Notebookを用いた演習を行う。</p> <p>・自然言語処理(3回) 自然言語を扱う技術について講述する。形態素解析などの文書解析基本技術、機械学習と大規模言語モデルにより機械翻訳や文書要約などのタスクに応用する技術についてPython, Jupyter Notebookを用いた演習を行う。</p>									
----- 情報分析・管理演習(2)へ続く -----									

## 情報分析・管理演習(2)

### 【履修要件】

- 情報分析・管理論を，原則として受講していること．
- プログラミングを伴う演習を受講する意欲があれば，プログラミング経験は問わないが，PCの基本的な使用方法を習得していること，何らかのソフトウェアをダウンロードした経験があることが望ましい．
- 各自のノートPC等の持ち込みを前提とする．

### 【成績評価の方法・観点】

演習で扱った情報分析・管理技術を理解し，具体的な問題に応用し，実際にその解を計算できるようになることが達成目標である．出席状況，ならびに毎週のレポートによってこの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】

(参考書)

- D. Easley, J. Kleinberg (著), 浅野孝夫, 浅野泰仁 (翻訳) 『ネットワーク・大衆・マーケット: 現代社会の複雑な連結性についての推論 Crowds, and Markets: Reasoning About a Highly Connected World』 (共立出版 Cambridge University Press)
- Jon Kleinberg, Eva Tardos (著), 浅野孝夫, 浅野泰仁, 小野孝男, 平田富夫 (翻訳) 『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)
- H. Garcia-Molina, J. D. Ullman, J. Widom 『Database Systems: The Complete Book』 (Prentice Hall)
- C.J. Date 『An Introduction to Database Systems』 (Addison Wesley)
- 吉川正俊 『データベースの基礎』 (オーム社)
- W. Bruce Croft, Donald Metzler, Trevor Strohman 『Search Engines: Information Retrieval in Practice』 (Addison-Wesley)
- Stefan Buettcher, Charles L. A. Clarke, Gordon V. Cormack 『Information Retrieval: Implementing and Evaluating Search Engines』 (The MIT Press)
- Jenifer Tidwell (著), ソシオメディア株式会社 (監訳), 浅野 紀予 (訳) 『デザイン・インタフェース ~ パターンによる実践的インタラクションデザイン ~』 (オライリー・ジャパン)
- Ben Fry (著), 増井 俊之 (監訳) (監修), 加藤 慶彦 (翻訳) 『ビジュアルライジング・データ Processingによる情報視覚化手法』 (オライリー・ジャパン)
- Marti Hearst 『Search User Interfaces』 (Cambridge University Press)

### 【授業外学修(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)】

- 前期・後期共に同一内容のリピート科目である．
- 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．
- 単位を必要としない学生の聴講は，受講希望人数などにより，受け入れられないこともある．
- 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．

情報分析・管理演習(3)へ続く

情報分析・管理演習(3)

東風上: kochigami@robot.soc.i.kyoto-u.ac.jp, Huang: huang@nlp.ist.i.kyoto-u.ac.jp