

科目ナンバリング		G-LAS12 80022 LJ13									
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 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	2単位	週コマ数	1コマ		授業形態	講義（対面授業科目）			
開講年度・ 開講期	2024・前期		曜時限	月4		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】											
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講述する。特に、具体的な問題に対して、それらの技術をどのように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、各自の分野において、応用できるようになることを目標とする。											
【到達目標】											
【研究科横断型教育の概要・目的】 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。											
【授業計画と内容】											
・ガイダンス (1回) 講義全体の概要											
・問題のモデル化と問題の解き方 (2回) 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、情報科学の他分野への応用について学ぶ。											
・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（アソシエーションルール、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）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。											
・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説明する。											
・自然言語処理（3回）											
情報分析・管理論(2)へ続く↓↓↓											

情報分析・管理論(2)

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を学ぶ。

【履修要件】

- 「情報分析・管理演習」を，原則として受講すること．
- 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】

講義で扱った情報分析・管理技術の仕組みを理解し，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．各回の出席状況，ならびに数回ごとに扱うテーマに基づいたレポートによって，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)

Jon M. Kleinberg 『Authoritative Sources in a Hyperlinked Environment』 (Journal of the ACM (JACM), 46(5), pages 604-632, 1999.)

L. Page, S. Brin, R. Motwani, and T. Winograd 『The PageRank Citation Ranking: Bringing Order to the Web』 (Technical Report SIDL-WP-1999-0120, Stanford Digital Library Technologies Project, 1998.)

【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．

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