

データマイニング -数理とPythonでデータサイエンス・AIを学ぶ-

単位数		ナンバリングコード	
2		DIF427	
	教員名	越野 一博	
	専門	医用画像工学	
	出身校等	北見工業大学	
	現職	北海道情報大学 経営情報学部 教授	
授業形態			
前期印刷授業・後期印刷授業			
授業範囲		試験範囲	
Moodle（通信教育部POLITE）に配置した授業資料の内容全て（補足資料は除く）		Moodle（通信教育部POLITE）に配置した授業資料の内容全て（補足資料は除く）	
		【試験時参照許可物】 一切自由 ※ただしWebページ（通信教育部POLITEを除く）と生成系AIの参照は不可とする。	
科目の概要			
生成・保存されるデジタルデータは、その量および種類ともに年を追うごとに増加しています。本科目では、大量のデータに対する分類、特徴の抽出や予測を可能とする機械学習を学びます。知識とあわせて、実習によるデータ処理の技術を習得します。 授業は、e-Learning システム（Moodle）を利用して行います。			
授業における学修の到達目標			
1. データマイニングの概念と役割を理解する 2. 目的に応じて統計的機械学習手法を選択し、データ分析が行える 3. 分析結果の数値やグラフの意味を説明できるようになる 4. 実習（プログラミング）をとおして、Pythonと機械学習用ライブラリの基本的な使い方を習得する			
講義の方針・計画			
第1回：データマイニングの概要，データの統計量と可視化 第2回：単回帰 第3回：重回帰，データの前処理 第4回：多項式回帰，汎化性能 第5回：非線形回帰 第6回：サポートベクトルマシン分類器，2クラス分類の性能指標 第7回：カーネルサポートベクトルマシン分類器，ハイパーパラメータのチューニング，多クラス分類の性能指標 第8回：決定木による分類 第9回：アンサンブル法による分類 第10回：決定木による回帰 第11回：アンサンブル法による回帰 第12回：K-means法，DBSCAN 第13回：ネットワーク分析，確率的クラスターリング			

講義の方針・計画
第14回：主成分分析 第15回：次元削減
準備学習
授業範囲の内容全体を2単位で90時間かけて学習することを目安としています。 和の記号、ベクトルと行列、微分について理解し計算できることが望ましいです。また、何らかのプログラミング言語を使った経験があると実習やプログラミング課題の助けとなります。
課題(試験やレポート等)に対するフィードバック方法
通信教育部POLITEや無限大キャンパスを利用して行います。
成績評価の方法およびその基準
Moodle試験：50% Moodle小テスト：50%
教科書
なし
参考書
Python 機械学習プログラミング 達人データサイエンティストによる理論と実践 第3版 Pythonデータサイエンスハンドブック Jupyter、NumPy、pandas、Matplotlib、scikit-learnを使ったデータ分析、機械学習
その他
なし
試験期間
シラバス検索画面トップページ（ https://syllabus-tsushin.do-johodai.ac.jp/ ）下部の「2025学年暦」を参照
学習プリント
あり
教職科目
関連受講科目
なし
担当教員の実務経験
2003年度から2018年度まで国立の研究所に所属し、医用画像の処理・解析研究を遂行した。その間、医師との共同研究も実施した。研究テーマの一つとして、画像から得られた複数の定量値や検査数値を対象として、診断や病態の進行、予後予測に有用な指標の探索を機械学習を利用して行ってきた。