

授業科目名	データサイエンス実践演習	科目コード	G4603P03
英文名	Practical Seminar on Data Science II		

科目区分	経営情報専攻科目 - データサイエンス		
------	---------------------	--	--

職名	准教授	担当教員名	河野 信
学部	現代社会学部	学科	現代社会学科
曜日	水曜日	時限	1限目
開講時期	3年前期	授業の方法	演習
必修・選択	選択	単位数	2単位

授業の概要	<p>本講義では実際のデータに触れながら、RやPythonといったデータ解析用のプログラミング言語を利用して、モデリングの手法や機械学習、ディープラーニングについて学ぶ。データの中から関連性を抽出し現象の解明や要因の分析に役立つ知識を得たり、データに潜む関連性をもとに予測を行う方法について学習する。本講義ではPCを利用する。「データサイエンス実践演習I」、「情報統計学I・II」、「プログラミングの基礎」を受講していることが望ましい。（担当教員：企業での実務経験有り）</p>			
-------	---	--	--	--

キーワード	R	Python	データ分析	モデリング	機械学習
-------	---	--------	-------	-------	------

到達目標	統計的なモデリングを実行するためのデータの準備と加工ができる（40%）				
	機械学習で予測モデルを作成できる（30%）				
	ディープラーニングで予測モデルを作成できる30%）				

ディプロマポリシー	3.スペシャリストとしての能力（専門性の向上）	
カリキュラムポリシー	専門性をさらに高める実習・演習	

キー・コンピテンシー（重視する能力）				
コミュニケーション力	協働力	課題解決力	人間理解力	教育支援力

教授方法（授業方法）					
知識教授型	対話型授業	演習・反復型授業	グループ演習	地域フィールドワーク	授業外学修指導・自主活動

授業計画

回数	授業内容 詳細	標準時間
第1回	ガイダンス	
	【予習】シラバスを読んでおくこと	30分
	【復習】ガイダンス資料を確認し、受講する場合には受講登録をすること	30分
第2回	モデリングの準備 - データの準備と加工、分析とモデリングの手法	
	【予習】教科書4.1 (p220 ~ 225) を読んでおく	60分
	【復習】作成・実行したプログラムの入力値を変えて実行し、結果を確認すること	60分
第3回	データの加工1 - データのクレンジング、カテゴリ変数の加工、数値変数の加工とスケールリング	
	【予習】教科書4.2.1 ~ 4.2.3 (p226 ~ 238) を読んでおく	60分
	【復習】作成・実行したプログラムの入力値を変えて実行し、結果を確認すること	60分
第4回	データの加工2 - 分布の形を変える - 対数変換とロジット変換	
	【予習】教科書4.2.4 (p239 ~ 252) を読んでおく	60分
	【復習】作成・実行したプログラムの入力値を変えて実行し、結果を確認すること	60分
第5回	データの加工3 - 欠損値の処理、外れ値の処理	
	【予習】教科書4.2.5 ~ 4.2.6 (p252 ~ 263) を読んでおく	60分
	【復習】	60分
第6回	モデリングの手法1 - グループに分ける - クラスタリング	
	【予習】教科書4.3.1 (p264 ~ 275) を読んでおく	60分
	【復習】作成・実行したプログラムの入力値を変えて実行し、結果を確認すること	60分
第7回	モデリングの手法2 - 指標を集約する - 因子分析と主成分分析	
	【予習】教科書4.3.2 (p275 ~ 302) を読んでおく	60分
	【復習】作成・実行したプログラムの入力値を変えて実行し、結果を確認すること	60分
第8回	モデリングの手法3 - 一般化線形モデル (GLM) とステップワイズ法、2値データを目的とする分析 - ロジスティック回帰	
	【予習】教科書4.3.3 ~ 4.3.4 (p302 ~ 321) を読んでおく	60分
	【復習】作成・実行したプログラムの入力値を変えて実行し、結果を確認すること	60分
第9回	モデリングの手法4 - セグメントの抽出とその特徴の分析 - 決定木	
	【予習】教科書4.3.5 (p321 ~ 331) を読んでおく	60分
	【復習】作成・実行したプログラムの入力値を変えて実行し、結果を確認すること	60分

	因果推論 - データから因果関係を明らかにする - 統計的因果推論、因果関係に基づく変数選択	
第10回	【予習】教科書4.4 (p332 ~ 346) を読んでおく	60分
	【復習】作成・実行したプログラムの入力値を変えて実行し、結果を確認すること	60分
	機械学習の目的と手順 - 機械学習の基本、機械学習の手順、データ準備に関わる問題、特徴抽出と特徴ベクトル	
第11回	【予習】教科書5.1 (p348 ~ 369) を読んでおく	60分
	【復習】作成・実行したプログラムの入力値を変えて実行し、結果を確認すること	60分
	機械学習の実行 - 機械学習ライブラリの活用 - scikit-learn、機械学習アルゴリズムの例 - ランダムフォレスト、サポートベクターマシン (SVM)、機械学習の実行例	
第12回	【予習】教科書5.2 (p370 ~ 389) を読んでおく	60分
	【復習】作成・実行したプログラムの入力値を変えて実行し、結果を確認すること	60分
	ディープラーニング - ニューラルネットワーク、ディープラーニングを支える技術、ディープラーニング・フレームワーク	
第13回	【予習】教科書5.3.1 ~ 5.3.3 (p390 ~ 398) を読んでおく	60分
	【復習】作成・実行したプログラムの入力値を変えて実行し、結果を確認すること	60分
	ディープラーニング - ディープラーニングの実行、生成モデル	
第14回	【予習】教科書5.3.4 ~ 5.3.5 (p399 ~ 412) を読んでおく	60分
	【復習】作成・実行したプログラムの入力値を変えて実行し、結果を確認すること	60分
	まとめと総復習	
第15回	【予習】これまで学んだことを復習し、わからないところを調べる	180分
	【復習】レポート課題を作成し、提出すること	180分
第16回	【予習】	
	【復習】	

評価方法	講義期間中の課題提出（50%）+ 期末レポート（50%） なお評価する基準は、「富山国際大学成績評価基準」にしたがって評価する。		
使用資料 <テキスト>	RとPythonで学ぶ実践的データサイエンス & 機械学習、有賀友紀・大橋俊介著、技術	使用資料 <参考図書>	データサイエンス・オンライン講座 社会 人のためのデータサイエンス入門（総務省
授業外学修等			
授業外質問方法	いつでもどうぞ		
オフィス・アワー	月曜5限		