

徳山工業高等専門学校		開講年度	平成31年度 (2019年度)	授業科目	自然言語処理
科目基礎情報					
科目番号	0024		科目区分	専門 / 選択	
授業形態	講義		単位の種別と単位数	学修単位: 2	
開設学科	情報電子工学専攻		対象学年	専2	
開設期	前期		週時間数	2	
教科書/教材	長尾真:「自然言語処理」, 岩波書店; 高村大也:「言語処理のための機械学習入門」, コロナ社				
担当教員	高山 泰博				
到達目標					
自然言語処理に関する用語の具体的内容、及びその有効性について、また、各処理段階における基本的な表現法やアルゴリズムについて説明・活用できる。					
ルーブリック					
	理想的な到達レベルの目安	標準的な到達レベルの目安	未到達レベルの目安		
総合的能力	自然言語処理の基本事項を理解して十分に説明・活用できる。	自然言語処理の基本事項を理解し説明できる。	自然言語処理の基本事項を理解できない。		
学科の到達目標項目との関係					
教育方法等					
概要	我々が日常使用する言葉(日本語や英語)は、プログラミング言語(C言語等)の人工言語と対比して、自然言語と呼ばれている。本講義では、自然言語処理の基本的段階である形態素解析、構文解析、意味解析に関する基本事項を学ぶ。また、現代的な統計的言語処理の基礎についても学ぶ。この科目は、企業で実際に自然言語処理応用システム開発に関わる実務を担当していた教員が、その経験を生かして、自然言語処理の基礎と応用についての講義形式および演習形式で授業を行うものである。				
授業の進め方・方法	各自の担当箇所を割当て、輪講形式で授業を進める。輪講で用いたレジュメはコメントを反映して次回の授業までに提出する。事前事後学習として、各回の授業の予習・復習内容を授業まとめノートの形で作成したり、アルゴリズム確認のためのプログラムを作成してレポートを作成したりを行う(計約60時間)。				
注意点	輪講形式で授業をすすめるため、事前の予習範囲を確実に予習して、その内容を説明できるようにしておくことが必須である。				
授業計画					
	週	授業内容	週ごとの到達目標		
前期	1stQ	1週	自然言語処理の概要 【事前事後学習の内容(4時間)】まとめノート作成	授業の進め方を説明し、自然言語処理とは何か、自然言語処理の歴史について学ぶ。	
	2週	統計的言語処理入門 【事前事後学習の内容(4時間)】まとめノート作成	自然言語の数学的表現方法について学び、naive Bayes法をもとに統計的な言語処理の考え方を習得する。		
	3週	言語の統計 【事前事後学習の内容(4時間)】まとめノート作成	言語の統計、統計の取り方、ジップの法則、マルコフ・モデル、シャノンの理論、テキスト間の類似性の検出などについて学ぶ。		
	4週	言語の有限オートマトンモデル 【事前事後学習の内容(4時間)】まとめノート作成	言語の基本的概念、言語の性質、有限オートマトン、正規言語、及びオートマトンと言語の関係などについて学ぶ。		
	5週	文脈自由型言語 【事前事後学習の内容(4時間)】まとめノート作成	形式文法、文脈自由型句構造言語、その性質、プッシュダウンオートマトン、オートマトンと言語の関係について学ぶ。		
	6週	記号論理と言語 【事前事後学習の内容(4時間)】まとめノート作成	記号論理、記号論理式の意味と文の意味について学ぶ。		
	7週	形態素と形態素解析(1) 【事前事後学習の内容(4時間)】まとめノート作成	形態素解析とは? 形態素や語の定義などについて学ぶ。		
	8週	形態素と形態素解析(2) 【事前事後学習の内容(8時間)】形態素解析プログラム作成	形態素解析のアルゴリズムについて学ぶ。		
	9週	構文解析(文脈自由文法に基づく構文解析I) 【事前事後学習の内容(4時間)】まとめノート作成	構文解析とは?。自然言語の文法をより自然に表現できるように前終端記号を導入した文脈自由文法を定義し、それに基づく構文解析アルゴリズムについて学ぶ。		
	10週	構文解析(文脈自由文法に基づく構文解析II) 【事前事後学習の内容(4時間)】まとめノート作成	文脈自由文法に基づく各種構文解析アルゴリズムについて学ぶ。		
	11週	構文解析(文脈自由文法の拡張) 【事前事後学習の内容(4時間)】まとめノート作成	実用規模の自然言語の文法を文脈自由文法として記述することは規則の数が膨大になり、文法の見通しを著しく低下させる。文法の見通しが悪くならないように文脈自由文法を拡張した文法記述法について学ぶ。		
	12週	構文解析(構文解析における優先規則) 【事前事後学習の内容(8時間)】構文解析プログラム作成	入力文の構造を一意に決定せず、可能な解に優先順位付けを行ったりして、ある程度の精度で一意的解を求める。そのときの優先順位について学ぶ。		
	13週	意味解析 【事前事後学習の内容(4時間)】まとめノート作成	文の意味解析とは何かという問題について学ぶ。意味的あいまい性の解消を選択制限、連想関係に基づく方法について学ぶ。		
	14週	自然言語処理の応用システム 【事前事後学習の内容(4時間)】まとめノート作成	自然言語処理技術を用いた応用システムの事例について学ぶ。		
	15週	試験	講義で学んだ内容の理解度を問う。		
	16週	まとめ 【事前事後学習の内容(3時間)】科目まとめノート作成	答案の返却、解説を行う。		
モデルコアカリキュラムの学習内容及到達目標					
分類	分野	学習内容	学習内容の到達目標	到達レベル	授業週

専門的能力	分野別の専門工学	情報系分野	システムプログラム	コンピュータシステムにおけるオペレーティングシステムの位置づけを説明できる。	4	前1,前14
				プロセス管理やスケジューリングなどCPUの仮想化について説明できる。	4	前1,前14
				排他制御の基本的な考え方について説明できる。	3	前1,前14
				記憶管理の基本的な考え方について説明できる。	3	前1,前14
				形式言語の概念について説明できる。	5	前4
				オートマトンの概念について説明できる。	5	前4,前5
				コンパイラの役割と仕組みについて説明できる。	5	前4
				正規表現と有限オートマトンの関係を説明できる。	5	前4
		その他の学習内容	少なくとも一つの具体的なコンピュータシステムについて、起動・終了やファイル操作など、基本的操作が行える。	5	前13,前14	
			少なくとも一つの具体的なオフィススイート等を使って、文書作成や図表作成ができ、報告書やプレゼンテーション資料を作成できる。	5	前13	
			少なくとも一つのメールツールとWebブラウザを使って、メールの送受信とWebブラウジングを行うことができる。	5	前13	
			コンピュータウイルスやフィッシングなど、コンピュータを扱っている際に遭遇しうる代表的な脅威について説明できる。	3	前14	
			コンピュータを扱っている際に遭遇しうる脅威に対する対策例について説明できる。	3	前14	
			データモデル、データベース設計法に関する基本的な概念を説明できる。	4	前14	
			データベース言語を用いて基本的なデータ問い合わせを記述できる。	4	前14	
メディア情報の主要な表現形式や処理技法について説明できる。	4	前13				
情報を離散化する際に必要な技術ならびに生じる現象について説明できる。	5	前2				

評価割合

	試験	課題レポート	発表レジュメ	態度			合計
総合評価割合	60	20	20	0	0	0	100
総合的能力	60	20	20	0	0	0	100