2019年度 授業シラバスの詳細内容

○基本情報

科目名(英)	自然言語処理及び演習（Natural Language Processingand Exercises ）		
ナンバリングコード	P20404	大分類 / 難易度 科目分野	情報メディア学科 専門科目 / 標準レベル メディア処理
単位数	3	配当学年 / 開講期	2 / 後期
必修・選択区分	コース必修：情報工学コース 選択：メディアデザインコース、こども・情報教育コース、情報コミュニケーションコース		
授業コード	P040451	クラス名	
担当教員名	赤星　哲也		
履修上の注意、 履修条件	本科目は講義とパソコン演習を組み合わせた授業形式を取っています。プログラミング演習にはPython言語を使用します(「プログラミング基礎(1年必修科目)」を履修済みであることが望ましい)。授業資料を配付するため、教科書は特に指定しませんが、授業資料、並びに、参考文献、指定図書、Webサイト等を用いて、予習・復習に取り組んでください。		
教科書	データにて提供、プリントを配付		
参考文献及び指定図書	備考欄に記載		
関連科目	情報メディア入門、プログラミング基礎、研究ゼミナールB、卒業研究		

○成績評価の指標 / ○成績評価基準（合計100点）

到達目標の観点	到達目標	テスト（期末試験・中間確認等）	提出物（レポート・作品等）	無形成果（発表・その他）
【関心・意欲・態度】	①授業に積極的に取り組む。 ②欠席、遅刻をしない。 ③授業時に課された課題を全て提出する。			15点
【知識・理解】	①要素技術（形態素解析、コスト最小法、コーパス、英語の品詞付与、構文解析、CKY法)について理解している。 ②応用技術（情報検索、機械翻訳、対話システム、情報抽出）について理解している。	65点		
【技能・表現・ コミュニケーション】				
【思考・判断・創造】	①学修した知識を応用して課題を解くことができる。 ②期限内に課題を完成させ提出する。		20点	

○到達目標に対する到達度の目安、および、成績評価の補足

課題毎に次の評価基準を用いて採点し、合計を20点満点に換算して評価を行います。
4点：課題内容の到達目標を満たしており、かつ、期限内に提出している。
3点：課題内容の到達目標を満たしているが、、期限内に提出できなかった。
2点：課題内容の到達目標にはやや不十分な内容だが、期限内に提出している。
1点：課題内容の到達目標にははやや不十分な内容であり、かつ提出内に提出できなかった。
0点：課題が提出されていない。
授業中の取り組み、欠席・遅刻、課題の提出状況を無形成果として15点満点で評価します。
授業内評価、中間評価の合計を65点満点に換算して評価を行います。

○授業の目的・概要等

授業の目的	本科目は、情報メディア学科情報工学コースのコース必修科目です。情報工学コースが扱うメディア処理分野における言語メディア処理領域をカバーしています。自然言語処理とはAI(人工知能)技術の一領域で、人間の「ことば」をコンピュータで扱うことを目的とした研究・技術分野を指します。本科目は自然言語処理に関する要素技術、並びに応用技術について学修します。具体的には、要素技術として形態素解析、コーパス、英語の品詞付与、構文解析、CKY法を、応用技術として情報検索、対話システム、機械翻訳を取り上げます。	
授業の概要	本科目は半期週2コマの授業です。講義形式の座学に加え、パソコン等を用いた演習形式を併用しています。演習課題は必ず期限内に完成させて提出してください。全30回の授業の内、前半と後半に分けて中間評価を各1回ずつ、さらに、随時、小テスト等による授業内評価を行います。授業の前半部分では形態素解析、コーパス、英語の品詞付与を扱います。授業の後半部分では構文解析、CKY法、情報検索、対話システム、機械翻訳を扱います。課題、授業内評価、中間評価については次回に解説を行い、学修した知識の定着をはかります。パソコン演習を通して、パソコンを用いテキスト情報を処理する技術を習得します。演習では単独で行うものと、グループワークで行うものとがあります。	
授業の運営方法	（1）授業の形式	「講義形式」
	（2）複数担当の場合の方式	「該当しない」
	（3）アクティブ・ラーニング	「グループワーク」
地域志向科目	該当しない	
実務経験のある教員による授業科目	企業の研究所にて、ハイパーテキストの研究開発に従事	

○備考欄

授業を理解する上で、以下の書籍やWebサイト、月刊誌(図書館データベースから閲覧可)を積極的に参考にしてください。

（書籍）
「働きたくないイタチと言葉がわかるロボット　人工知能から考える「人と言葉」」 朝日出版社
「自然言語処理」 放送大学教材
「自然言語処理概論」 サイエンス社
「はじめての自然言語処理」 森北出版
「言語研究のためのプログラミング入門 ：Pythonを活用したテキスト処理」 開拓社
「Pythonで動かして学ぶ 自然言語処理入門」 翔泳社
「自然言語処理の基礎」 コロナ社
「自然言語処理の基本と技術」 翔泳社
「自然言語処理（岩波講座ソフトウェア科学15）」 岩波書店
「言語研究のための統計入門」 くろしお出版
「入門　自然言語処理」オライリージャパン
「IT Text　自然言語処理」オーム社
「言語処理のための機械学習入門」　コロナ社,
「言語処理学事典」 共立出版

（Webサイト）
言語情報処理ポータル　hhttp://www.jaist.ac.jp/project/NLP_Portal/
言語グリッドLanguage Grid Playgraound　https://langrid.org/playground/

（月刊誌：図書館データベースから閲覧可）
「日経ソフトウェア」
※NBU大学図書館・情報検索サービス「日経BP社記事検索サービス」を使って閲覧することができます。
http://bizboard.nikkeibp.co.jp/daigaku/

2019年度 授業シラバスの詳細内容

○授業計画　科 目 名 : 自然言語処理及び演習（Natural Language Processingand Exe　授業コード:P040451
担当教員: 赤星　哲也

学修内容

1. ガイダンス、テキスト処理
この科目の目的、到達目標、学修内容、学修方法（受講心得）、成績評価方法について説明します。自然言語処理を用いた様々なデモを通して、自然言語処理の技術を概観してみます。また、テキストエディタを用いたテキスト処理の技法を習得します。

予習: 本授業シラバスの確認、学生便覧掲載の「専門教育科目連携表」の確認　（約2.5h）
復習: 配付資料、並びに、参考文献・指定図書による復習　（約2.5h）

2. 日本語文法、テキストマイニング
日本語文法について学修します。テキストマイニングの演習を通して、日本語文法、特に、品詞に関する知識の理解を深めます。

予習: 配付資料、並びに、参考文献・指定図書による予習　（約2.5h）
復習: 配付資料、並びに、参考文献・指定図書による復習　（約2.5h）

3. 形態素解析:基本アルゴリズム
形態素解析の基本的なアルゴリズムを学修します。形態素解析とは、①文を単語(形態素)に切り分け、②単語の品詞を決定し、③単語を原形に戻す技術をさします。日本語の形態素解析では、①の単語の切り分けが最も重要な処理となります。

予習: 配付資料、並びに、参考文献・指定図書による予習　（約2.5h）
復習: 配付資料、並びに、参考文献・指定図書による復習　（約2.5h）

4. 形態素解析:コスト最小法　コーパス
形態素解析の代表的なアルゴリズムの一つ、コスト最小法について学修します。コスト最小法とは、単語(形態素)にコストを付与し、文を構成する全単語のコストの合計が最小となる解析結果を最適解とするアルゴリズムです。また、コーパスについて学修します。コーパスとは自然言語処理に必要となる知識の獲得を目的として集められたテキストデータをさします。具体的なコーパスとして、BCCW、TWC等を取り上げ、コーパスとはどのようなものかを確認してみます。

予習: 配付資料、並びに、参考文献・指定図書による予習　（約2.5h）
復習: 配付資料、並びに、参考文献・指定図書による復習　（約2.5h）

5. コーパス
前回に引き続き、コーパスについて学修します。具体的には、マルコフモデル、1階マルコフモデル、n-gram言語モデル、bigram言語モデル、条件付き確率、bigramの出現確率、コーパスを用いた文の出現確率の求め方を扱います。また、機械学習の一つ、ナイーブベイズ分類器を用いた分類問題について学修します。

予習: 配付資料、並びに、参考文献・指定図書による予習　（約2.5h）
復習: 配付資料、並びに、参考文献・指定図書による復習　（約2.5h）

6. 英語の品詞付与
英語の形態素解析に必要となる品詞付与のアルゴリズムについて学修します。英語の形態素解析では文を構成する単語の品詞を決めることが最も重要な処理となります。授業では隠れマルコフモデルを用いた統計的言語モデル・アプローチのアルゴリズムを取り上げます。

予習: 配付資料、並びに、参考文献・指定図書による予習　（約3.5h）
復習: 配付資料、並びに、参考文献・指定図書による復習　（約1.5h）

7. 中間評価(1)、パソコン演習
これまでに学修した内容(形態素解析、コスト最小法、コーパス、英語の品詞付与)が理解できているかを確認するための中間評価を行います。また、自然言語処理の技術を用いたパソコン演習に取り組むことで、自然言語処理の技術に関する理解を深めます。

予習: 中間評価(1)の試験勉強　配付資料、並びに、参考文献・指定図書による予習　（約2.5h）
復習: 配付資料、並びに、参考文献・指定図書による復習　課題の提出　（約2.5h）

8. パソコン演習
自然言語処理の技術を用いたパソコン演習に取り組むことで、自然言語処理の技術に関する理解を深めます。

予習: 配付資料、並びに、参考文献・指定図書による予習　（約2.5h）
復習: 配付資料、並びに、参考文献・指定図書による復習　課題の提出　（約2.5h）

○授業計画　科 目 名 : 自然言語処理及び演習（Natural Language Processingand Exe　授業コード:P040451
担当教員: 赤星　哲也

学修内容

9. 構文解析:基本アルゴリズム
構文解析の基本的なアルゴリズムを学修します。構文解析とは文の構造を求める技術をさします。文の構造とは、たとえば、日本語では文節間の係り受け関係のこと、英語の場合ではSVOC等の文型のことです。授業では英語文を扱います。文法規則として文脈自由文法のチョムスキー標準形を用いた、最左導出のアルゴリズムを取り上げます。

予習: 配付資料、並びに、参考文献・指定図書による予習　（約2.5h）
復習: 配付資料、並びに、参考文献・指定図書による復習　（約2.5h）

10. 構文解析:基本アルゴリズム、構文解析:CKY法
前回に引き続き、構文解析の基本的なアルゴリズムを学修します。さらに、構文解析の代表的なアルゴリズムの一つ、CKY法について学修します。

予習: 配付資料、並びに、参考文献・指定図書による予習　（約2.5h）
復習: 配付資料、並びに、参考文献・指定図書による復習　（約2.5h）

11. 構文解析:CKY法
前回に引き続き、構文解析の代表的なアルゴリズムの一つ、CKY法について学修します。CKY法の理解を深めるために、Pythonによる簡単なCKY法シミュレータを作成し、アルゴルズムの様子を視覚的に確認していきます。

予習: 配付資料、並びに、参考文献・指定図書による予習　（約2.5h）
復習: 配付資料、並びに、参考文献・指定図書による復習　（約2.5h）

12. 情報検索:検索モデル
自然言語処理の応用分野として情報検索があります。情報検索の検索モデルについて学修します。具体的には、転置インデックス、TF-IDF法、ベクトル空間モデル、コサイン類似度、ランキングを扱います。理解を深めるために、Excelを用いたパソコン演習を行います。

予習: 配付資料、並びに、参考文献・指定図書による予習　（約4.0h）
復習: 配付資料、並びに、参考文献・指定図書による復習　（約1.0h）

13. 情報検索:検索モデル、情報検索:評価方法
前回に引き続き、情報検索の検索モデルについて学修します。さらに、情報検索システムの評価方法について学修します。具体的には、適合率、再現率、F値、平均適合率、ページランクを扱います。

予習: 配付資料、並びに、参考文献・指定図書による予習　（約.0h）
復習: 配付資料、並びに、参考文献・指定図書による復習　（約5.0h）

14. 対話システム・機械翻訳、中間評価(2)
自然言語処理の応用分野である対話システム、機械翻訳について学修します。また、中間評価(1)以降の内容が理解できているかを確認するための中間評価を行います。

予習: 配付資料、並びに、参考文献・指定図書による予習　中間評価(2)の試験勉強　（約.0h）
復習: 配付資料、並びに、参考文献・指定図書による復習　（約5.0h）

15. パソコン演習
自然言語処理の技術を用いたパソコン演習に取り組むことで、自然言語処理の技術に関する理解を深めます。

予習: 配付資料、並びに、参考文献・指定図書による予習　（約.0h）
復習: 配付資料、並びに、参考文献・指定図書による復習　課題の提出　（約5.0h）

16.