

仙台高等専門学校		開講年度	平成31年度 (2019年度)	授業科目	分散コンピューティング	
<b>科目基礎情報</b>						
科目番号	0261		科目区分	専門 / 選択		
授業形態	授業		単位の種別と単位数	学修単位: 2		
開設学科	情報ネットワーク工学科		対象学年	5		
開設期	後期		週時間数	2		
教科書/教材	「パターンでわかるHadoop MapReduce」三木大知著 (翔泳社)					
担当教員	千葉 慎二					
<b>到達目標</b>						
<p>情報通信ネットワークを利用したアプリケーションの作成方法を理解している。  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理解している。  処理形態の面でのコンピュータシステムの種類である集中処理システムと分散処理システムについて, それぞれの特徴と代表的な例を説明できる。</p>						
<b>ルーブリック</b>						
		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		
ビッグデータと分散処理の概要		分散処理に必要な技術を理解し, HDFSやMapReduceの仕組みを概ね説明できている。	Hadoopの関連技術について理解し, その役割を概ね説明できている。	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理解できていない。		
分散処理の仕組みの理解		ある程度複雑な処理について, 分散処理の流れを概ね説明できる。	簡単な処理パターンについて処理記述のポイントを概ね説明できている。	MapReduceの各フェーズの内容や処理の流れを理解できていない。		
分散処理の応用プログラミングに関する理解		分散処理の効率を考慮したPigプログラミングができ, 課題解決への応用が概ねできている。	Pigの記述と処理パターンとの対応を理解できている。	Pigの記述について理解できていない。		
<b>学科の到達目標項目との関係</b>						
学習・教育到達度目標 1 通信機器や情報通信システム構築に必要なハードウェア・ソフトウェアの知識と技術の習得						
<b>教育方法等</b>						
概要	5学年前期の「分散コンピューティング I」で学んだ分散コンピューティングの基礎概念に続き, 本科目では分散コンピューティングの応用について学習する。ネットショップの購入履歴, Twitterのつぶやきなど, 近年膨大なデータ (ビッグデータ) を解析することで様々な有益な情報を見つけ出すデータマイニングが行われている。ビッグデータを高速処理するためのコンピューティング環境の一つであるHadoopは, 様々な分散処理を単純な処理の組み合わせで実現できるオープンソースである。本科目ではHadoopによる分散処理の基礎技術を習得し, 株価チャートやアクセスログの解析等の現実的な問題への適用について学習する。					
授業の進め方・方法	授業内で実施する演習・実習 (100%) で評価する。					
注意点	本科目は, 分散コンピューティング I, ネットワークアーキテクチャ, ネットワークシステム開発科目に関連する。本授業は座学での講義とそれに関連する演習・実習を適時併用しながら進めていく。Hadoopでの処理の記述方法を理解するだけではなく, 現実のビッグデータに適用してデータマイニングできる応用力を身に付けることが大事であるので, 演習・実習を積極的に行い, 様々なデータを処理してみること。本科目で使用するHadoopの実習環境は無償で準備できるので, 自宅等で実習環境を用意できる者は予習復習に活用することが望ましい。					
<b>授業計画</b>						
		週	授業内容	週ごとの到達目標		
後期	3rdQ	1週	ビッグデータとは何か	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理解している。		
		2週	Hadoopのエコシステム	Hadoopの関連技術の概要が理解できる。		
		3週	分散ファイルシステム	HDFS(Hadoop Distributed File System)について理解できる。		
		4週	MapReduceの基本概念 I	各フェーズの役割を理解できる。		
		5週	MapReduceの基本概念 II	処理形態の面でのコンピュータシステムの種類である集中処理システムと分散処理システムについて, それぞれの特徴と代表的な例を説明できる。		
		6週	MapReduceの処理パターン I	処理パターンごとの処理記述のポイントを理解できる。		
		7週	MapReduceの処理パターン II	処理パターンごとの処理記述のポイントを理解できる。		
		8週	株価チャートの処理 I	Pigの使用方法を理解できる。		
	4thQ	9週	株価チャートの処理 II	適切な処理パターンで記述できる。		
		10週	テキスト解析	非構造化データ処理の記述について理解できる。		
		11週	アクセスログの分析 I	ユーザ定義関数を記述できる。		
		12週	アクセスログの分析 II	正規表現によるパターンマッチングを使用できる。		
		13週	鉄道情報の分析 I	データの結合について理解できる。		
		14週	鉄道情報の分析 II	情報通信ネットワークを利用したアプリケーションの作成方法を理解している。		
		15週	応用課題	オープンデータを用いた解析プログラムを開発できる。		
		16週				
<b>モデルコアカリキュラムの学習内容と到達目標</b>						
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	プログラミング	要求仕様に従って, いずれかの手法により動作するプログラムを設計することができる。	4	後15
				要求仕様に従って, いずれかの手法により動作するプログラムを実装することができる。	4	後15

			ソフトウェア	アルゴリズムの概念を説明できる。	3	後6,後7,後8
				与えられたアルゴリズムが問題を解決していく過程を説明できる。	3	後6,後7,後8
				同一の問題に対し、それを解決できる複数のアルゴリズムが存在しうることを説明できる。	3	後8,後9,後10,後12
				整列、探索など、基本的なアルゴリズムについて説明できる。	3	後8,後9,後10,後12
				コンピュータ内部でデータを表現する方法(データ構造)にはバリエーションがあることを説明できる。	4	後3,後4,後11
				同一の問題に対し、選択したデータ構造によってアルゴリズムが変化しうることを説明できる。	3	後4
				リスト構造、スタック、キュー、木構造などの基本的なデータ構造の概念と操作を説明できる。	4	後3,後4,後11
				ソフトウェアを中心としたシステム開発のプロセスを説明できる。	3	
			コンピュータシステム	集中処理システムについて、それぞれの特徴と代表的な例を説明できる。	4	後5
				分散処理システムについて、特徴と代表的な例を説明できる。	4	後5
			情報通信ネットワーク	情報通信ネットワークを利用したアプリケーションの作成方法を説明できる。	4	後14
			その他の学習内容	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	後1
			分野別の工学実験・実習能力	情報系分野【実験・実習能力】	情報系【実験・実習】	要求仕様にあったソフトウェア(アプリケーション)を構築するために必要なツールや開発環境を構築することができる。

評価割合

	試験	実習	相互評価	態度	ポートフォリオ	その他	合計
総合評価割合	0	100	0	0	0	0	100
基礎的能力	0	50	0	0	0	0	50
専門的能力	0	50	0	0	0	0	50
分野横断的能力	0	0	0	0	0	0	0