

教科目名 データマイニング (Data Mining)

学科名・学年 : 情報工学科 5 年 (教育プログラム 第 2 学年 ○科目)

単位数など : 選択 1 単位 (後期 1 コマ, 授業時間 23.25 時間)

担当教員 : 行天 啓二

授業の概要			
データマイニングとは、データの中から有用な情報を効率的に見つけるための技術の総称である。本講義では、はじめに、データマイニングに用いられる確率や、情報理論の基礎について学習し、データマイニングの基本的な手法を学習する。			
達成目標と評価方法		大分高専目標(B2), JABEE 目標(2.1②)	
(1) データマイニングに用いる確率, 情報理論, 統計について理解し, それを説明できる。(小テスト, 定期試験)			
(2) データマイニングの基本的な手法について理解し, それを説明できる。(小テスト, 定期試験)			
(3) 教師あり学習および教師なし学習によるデータマイニングについて理解し, それを説明できる。(小テスト, 定期試験)			
(4) テキストマイニングの概念および各種手法について理解し, それを説明できる。(小テスト, 定期試験)			
回	授 業 項 目	内 容	理解度の自己点検
1	データマイニングの概要	○データマイニングの基本的な考え方を理解できる。	【理解の度合い】
2	教師あり学習, 教師なし学習	○教師あり学習, 教師なし学習の違いを理解できる。	
3	分布, 頻出, 相関	○分布, 頻出, 相関に関するデータ解析方法について, 数式に基づいて説明することができる。	
4	決定木	○決定木の概念を理解し, 決定木を用いてデータを解析することができる。	
5	ベイジアンアルゴリズム	○ベイジアンアルゴリズムについて, 数式に基づいて説明することができる。	
6	回帰	○回帰について, 数式に基づいて説明することができる。	
7	隠れマルコフモデル	○隠れマルコフモデルについて, その概念を理解し, 各パラメータの意味や学習アルゴリズムについて説明することができる。	
8	後期中間試験		【試験の点数】 点
9	後期中間試験の解答と解説	○分からなかった部分を把握し, 理解できる。	【理解の度合い】
10	最近傍決定則	○最近傍決定則の概念を理解し, 処理手順を説明することができる。	
11	K-means 法	○K-means 法の概念を理解し, 処理手順を説明することができる。	
12	サポートベクトルマシン	○サポートベクトルマシンの概念を理解し, 処理手順を説明することができる。	
13	テキストマイニングの概要, 背景知識	○テキストマイニングの概要を理解し, 背景知識の内容およびその必要性について理解できる。	
14	テキスト分類	○テキスト分類の概念を理解し, 各手法について説明できる。	
15	後期期末試験		【試験の点数】 点
	後期期末試験の解答と解説		
履 修 上 の 注 意	確率, 情報理論, 線形代数の復習をすること。		【総合達成度】
教 科 書	なし		
参 考 図 書	Satnam Alag, 「集合知イン・アクション」, SoftBank Creative Ronan Feldman 他, 「テキストマイニングハンドブック」, 東京電機大学出版局		
自 学 上 の 注 意	確率, 情報理論, 線形代数の復習をすること。		
関 連 科 目	情報理論, 自然言語処理		
総 合 評 価	達成目標の(1)~(4)について, 定期試験, レポートで評価する。 総合評価 = 定期試験 * 0.7 + レポート * 0.3 再試験は実施しない。		【総合評価】 点