

電気通信大学 平成20年度シラバス

授業科目名	データマイニング		
英文授業科目名			
開講年度	2008年度	開講年次	3年次
開講学期	後学期	開講コース・課程	昼間コース
授業の方法	講義	単位数	2
科目区分	専門科目-学科専門科目-選択科目		
開講学科・専攻	システム工学科 人間コミュニケーション学科		
担当教官名	山本 渉		
居室	西5-601		

公開E-Mail	授業関連Webページ
授業関連Webページに記載(右のURLは学内限定です)	http://port.se.uec.ac.jp/incampus/moin.cgi/DataMiningIntroduction/2008

【主題および達成目標】

データマイニング(data mining)とは、金鉱が埋まっているデータの山を掘る(mine)、という意味の用語で、大規模データからの問題発見の手法の総称の一つです。データマイニングする、という横文字で何だか新しい雰囲気を出しそうですが、実際には、多くのデータを集めてくる、データを加工して分析ツールで扱える形に整える、金鉱を見つけるのに必要な情報をデータに追加する、などデータマイニングを実施している時間の多くは、データマイニングツール利用の準備に費やされます。

【前もって履修しておくべき科目】

コンピュータ・リテラシー、基礎プログラミング、情報リテラシー、情報リテラシー演習

【前もって履修しておくことが望ましい科目】

確率統計学第二、多変量解析などの講義を受講しておいていただくと、講義で使用する言葉にあらかじめ馴染みをもっておくことができます。他には特にありませんが、記述統計についての一定の知識は有しておいて下さると助かります。

【教科書等】

教科書は特に指定しません。
参考書は講義中に提示します。

【授業内容とその進め方】

下記の内容を、授業の進度に応じて、調整したり、内容を選別したりしながら、進めていく予定です。

概論:データマイニングとは(概論, レポート課題1)
相関1:多次元データのまとめ方
相関2:分割表の推論
相関3:OLAPとアソシエーション・ルール(連関分析)
相関4:演習(レポート課題2)
分類1:階層型クラスタリング
分類2:非階層型クラスタリング
分類3:演習(レポート課題3)
判別1:線形判別分析
判別2:二次判別分析
判別3:線型モデルによる判別(ロジスティック判別1)
判別4:一般化加法モデルによる判別(ロジスティック判別2)
判別5:決定木
判別6:ニューラルネットワーク
判別7:サポートベクターマシン
判別8:演習(レポート課題4)
予測1:線形モデルと非線形モデル
予測2:一般化線形モデルによる回帰分析
予測3:一般化加法モデルによる回帰分析
予測4:その他の回帰分析
予測5:演習(レポート課題5)
補論1:データ・クレンジング
補論2:データ・レポジット

上記以外にも、必要に応じて、項目を追加することがあります。

例年、3回から4回の演習をシステム工学科計算機室で行い、講義の第一回と合わせて、4回から5回のレポート課題に取り組んでもらっています。
今年度から、レポートを早めに評価して返却することを始めてみようかと、考えています。

【成績評価方法及び評価基準(最低達成基準を含む)】

レポート課題と、演習の回の出席、を基礎資料とします。
原則として、それらの課題の難易度による重み付き和で評価をしています。

電気通信大学 平成20年度シラバス

【オフィスアワー：授業相談】

特に設けていません。
質問等は講義の後に尋ねて頂くか、
アポイントメントをとってから訪ねてください。

【学生へのメッセージ】

四年目の今年は、そろそろ内容が落ち着きそうです。

【その他】

本講義は、システム工学科計算機室を使用します。
毎年度の初頭に、アカウント登録のためのガイダンスが実施されており、
それに必ず参加してアカウントを作成・更新しておかないと、
この講義の単位を修得できません。
4月当初の掲示によく注意してください。