

電気通信大学 平成20年度シラバス

授業科目名	データマイニング特論		
英文授業科目名	Topics in Data Mining		
開講年度	2008年度	開講年次	
開講学期	後学期	開講コース・課程	博士前期課程
授業の方法	講義	単位数	2
科目区分	電気通信学研究科-システム工学専攻-専門科目		
開講学科・専攻	システム工学専攻		
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/DataMiningTopics/2008

<p>【主題および達成目標】</p> <p>データマイニングに現れる種々の手法を、基本的な事項から始めて、手法の実装に必要な理論を解説した後、実際の問題に適用するまでの一連の過程を通じて、学んでもらいます。毎年、次のようなトピックの中から、一つないし複数を選んで、講義します。</p> <p>Tree-based models and their enhancements (CART, boosting, bagging, etc.) Kernel-based machines (support vector machine) Network models (graphical modelling, Bayesian network, etc.)</p> <p>2004年度は、ベイズ流の統計的学習理論の中から特に、ベイジアン・ネットワークを主要なトピックとして扱いました。</p> <p>2005年度は、前半はSubsampling、後半はBayesian Networkを扱いました。 Subsamplingは、大量のデータがあるときにそれらを用いて、特定の確率分布を仮定せずに統計的な推測論を実施するアプローチの一つです。前半はこの技術の基本的なことを紹介した。後半のBayesian Networkは近年注目されている知識のネットワーク表現モデルの一つで、連続データか離散データの場合に限って紹介することになると思います。)</p> <p>2006年度は、ベイズ流の推論の基礎として、ベイズ統計学を扱いました。</p> <p>2007年は、"The Elements of Statistical Learning: Data Mining, Inference, and Prediction"をベースに、統計的学習理論について、講義しました。</p> <p>2008年度は、</p> <ul style="list-style-type: none"> ・ 統計的学習理論、特にアンサンブル学習を中心に。 ・ 大量なログデータの分析のための計数過程に基づく分析 ・ ベイズ統計学 <p>のどれかを講義する予定です。初回に来た方々と相談します。</p>

電気通信大学 平成20年度シラバス

講義の準備状況に応じて、このシラバスは更新されます。

2008年10月追記：相談の結果、今年度はベイズ統計学について、初歩から講義することにしました。

【前もって履修しておくべき科目】

なし。

【前もって履修しておくことが望ましい科目】

特にありませんが、確率・統計に関する基本的な知識は得ていることを前提に、講義を進めることがあります。

【教科書等】

教科書は特にありません。参考書は講義中に適宜、紹介します。

2008年10月追記：今年度の講義は

G. A. Young and R. L. Smith (2005)

"Essentials of Statistical Inference"

Cambridge Series in Statistical and Probabilistic Mathematics (No. 16)

を参考にしながら、進めて行きます。

【授業内容とその進め方】

資料を配布することもありますし、講義で使用する図表は予め用意しておくこともあります。

ただ、ノートを取ることは知識を吸収するための一手段として未だに有効、

と考えているので基本的には板書しながら講義を進める回が多いです。

【成績評価方法及び評価基準(最低達成基準を含む)】

講義時間中にプレゼンテーションを課したり、課題を示してそれに取り組んだレポートを提出してもらうなど、

講義内容に関連したタスクの成果、

あるいは学生個々の理解度に応じて、総合的に評価します。

【オフィスアワー：授業相談】

質問などは随時受けつけます。講義中の質問も歓迎します。

それ以外の時間については、アポイントメントをとってください。

【学生へのメッセージ】

データマイニングは様々な分野の技術を柔軟に応用する分野です。講義ではデータを扱う作業全般の中で、特にマイニングのステージだけにスポットを当てますが、データのクレンジングのステージもデータウェアハウスにデータを流し込む併合のステージも、それぞれ興味深い技術が用いられます。

データマイニング全般をカバーはせず、トピックを掘り下げることにはなりますが、興味を持って臨んでもらえればと思います。

【その他】

講義についてのメモは授業関連Webページを参照してください。