

# 確率解析学特論

(Advanced probability and data analysis)

種別・単位：講義・2単位（週1講時）

開講期：第2学期

担当者：工藤 峰一（コンピュータサイエンス専攻・数理計算科学講座・内線6852）

今井 英幸（コンピュータサイエンス専攻・数理計算科学講座・内線6856）

## 主題と目標

例からの学習は機械学習やパターン認識、データ解析において中核をなす概念である。その概念にあって、前期開講の「パターン情報処理特論」では主に線形モデルに関して深く議論した。本講義では、非線形モデルに力点を置いて各種の方法論を学ぶ。取り上げる話題は多岐に渡るものの、最新的话题を含めて伝える。受講者がこの分野での研究を自らの力で進められる力がつくよう、理解できる講義を心がける。

## 授業計画（項目、授業実施回数、内容）

項目	回数	内容
カーネル法	4	サポートベクターマシンを代表に近年脚光を浴びていカーネル法に関して、最新の研究動向を踏まえ、学習の枠組みで各種の方法論を眺める。
クラスタリング	3	教師なし学習としても知られるクラスタリングの各種方法を学ぶとともに得失を評価する。
アンサンブル法	2	複数の識別子を統合することでより高性能な識別を行う試みを紹介する。
ブースティングとバギング	3	アンサンブルの代表的な方法であるブースティングに関して検討する。また、バギングやランダムサブスペース法なども比較する。
発見的学習方法	3	プロトタイプ法、最近隣法、決定木など、統計的には性質がわかりにくいものの、各種の応用で有効な手法を概観する。

## 評価・教材・受講条件等

《評価》 講義内容に関する理解度を見る数回のレポートならびに各専門分野への応用可能性を論じたレポートによって学習の達成度を評価する。

《教材》 T. Hastie, R. Tibshirani and J. Friedman 著「The Elements of Statistical Learning」 2<sup>nd</sup> ed. (Springer, 2009) を用いる。また、適宜他の参考書および参考論文を示す。

《受講条件等》 基礎として学部レベルの線形代数および確率・統計の知識が必要であり、（多次元）確率変数の扱いに馴染んでいると理解しやすい。また、前期開講の「パターン情報処理特論」の継続内容とも言えるため、一緒に履修することでより理解が深まる。

キーワード：統計的学習、モデル選択、識別、クラスタリング、データマイニング