

コミュニケーション学研究室

(NTTコミュニケーション科学基礎研究所)

http://www.kecl.ntt.co.jp/rps/index.html



(写真左から)

客員教授：岩田 具治 tomoharu.iwata.gy@hco.ntt.co.jp

客員准教授：田中 佑典 yusuke.tanaka.rh@hco.ntt.co.jp

データを通じて実世界を解釈するための数理モデルとアルゴリズムを創出しよう

研究を始めるのに必要な知識・能力

機械学習に関連するものとして、確率統計、線形代数、微分に関する知識と、考案したアルゴリズムを実装するためのプログラミング、ツール活用能力が望まれます。

研究室の指導方針

入学後の修士1年生の前半では、授業で基礎的な学力を身につけ、上記記載の必要な知識・能力を強化すると共に、実データに既存の機械学習手法を適用して技術を活用する能力を身につけます。その後、オリジナル性のある研究テーマを設定し、モデル考案→アルゴリズムの実装→実データによる検証のサイクルを繰り返して、成果の完成度を高めていきます。最終的には、世界的に注目される主要国際会議での論文採録を目指します。

この研究で身につく能力

- ・近年幅広く用いられるようになった機械学習やディープラーニングを使いこなす能力。
- ・多種多様なデータを分析し、自分自身および第三者がそのデータの背後で起こった状況を把握できるようにする能力。
- ・世の中の事象を統計的に捉え、不確実性やランダム性がある中でも、将来の状況を予測したり、適切な施策を考案し実施する能力。
- ・機械学習やデータ分析手法を発展させ、新たな技術を開発したり、新領域に適用する能力。

修了生の活躍の場

情報通信産業、インターネットサービス産業、大学教職員、など

研究内容

インターネット上にある、あるいは実世界からセンサを通じて得られる多種多様なデータから、有益な情報を知識として抽出する機械学習およびデータマイニング技術に関する基礎研究を行っています。

今日我々が入手できるデータの量は非常に膨大で、またノイズも多く、人間が通り目を通して内容を理解し解釈する、という限界をはるかに超えています。従って、そのような膨大なデータを自動的に解析・解釈する方法が必要とされています。具体的には、以下のテーマに取り組んでいます。

・潜在トピック抽出技術

多様な大規模データに内在する隠れた構造(トピック)を抽出するための手法であるトピックモデルとその拡張の研究を行っています。トピックモデルは文書が生成される過程を表現した確率モデルです。トピックモデルを拡張し、可視化や推薦システム、購買ログ解析などへの適用を進めています。またディープラーニング、カーネル法を組み込むことにより、複雑なデータから隠れた構造を抽出することにも取り組んでいます。



・時間変化する関係データの解析・クラスタリング技術

ソーシャルネットワーク上の友人関係や企業間の取引関係など、時間とともに変化する「関係」の情報を対象として、関係ネットワークに潜むコミュニティクラスターの発見や関係の時間変化を追跡する技術を行っています。また、ノイズが多く含まれる関係データであっても有用なクラスターを発見するための研究も行っています。



・ソーシャルネットワーク上の情報伝播解析

ニュースや噂、評判情報はソーシャルネットワークを介して伝播します。この現象を確率モデルで表現し、将来のトレンド予測や潜在的な影響関係の推定の研究を行っています。



・時空間データ解析

センサ技術の発展とともに、気象、交通など様々な時空間データが容易に得られるようになってきています。非負値行列分解、ディープラーニング、ガウス過程などの技術を活用し、時間的・空間的に変化するデータから将来を予測したり、価値ある情報を抽出するための研究を行っています。



・センサを用いた実世界マイニング

加速度センサ、照度センサ、GPS、カメラ、マイクロホンなど様々なセンシングデバイスを用いて実世界の状況を観測し解釈して、人間にとって分かりやすい形式で提示するなどのセンサ情報処理に関する研究を行っています。

研究設備

計算機サーバ(マルチコアCPU、GPGPUボード搭載)

研究業績・共同研究・社会活動・外部資金など

■岩田具治, 2015. トピックモデル(機械学習プロフェッショナルシリーズ).

■Yoshikawa, Y., Iwata, T. and Sawada, H., 2014. Latent support measure machines for bag-of-words data classification. In Advances in Neural Information Processing Systems (pp. 1961-1969).

■Iwata, T., Shah, A. and Ghahramani, Z., 2013. August. Discovering latent influence in online social activities via shared cascade poisson processes. In Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 266-274). ACM.

■Sawada, H., Kameoka, H., Araki, S. and Ueda, N., 2013. Multichannel extensions of non-negative matrix factorization with complex-valued data. IEEE Transactions on Audio, Speech, and Language Processing, 21(6), pp.971-982.