

会 議 報 告

The 21st International Conference on Machine Learning (ICML) 2004 参加報告

1. ICML 2004 について

今年で 21 回目を迎える機械学習分野における最大の国際会議である International Conference on Machine Learning (ICML) は、2004 年 7 月 5 日から 8 日の 4 日間に渡り、カナダのアルバータ州バンフの Banff Park Lodge にて開催された。バンフはカナダの自然に囲まれた美しい町で、夏には全国より避暑に訪れる。日本でいうなら軽井沢といったところであろうか。

今回の ICML は、より理論に重点をおいた COLT (Annual Conference on Learning Theory) と、確率・統計に根ざした UAI (Uncertainty in Artificial Intelligence) との共同開催であり、大きな盛り上がりを見せた。会議の全体的な流れとしては、1 日目 (COLT の最終日で ICML の初日) にチュートリアルと、“Kernel Day” と名づけられた、COLT と ICML でのカーネル法に関する発表を纏めたセッションが開催された。2 日目以降は、招待講演、受賞した論文の講演に続き 3 つのセッションに分かれて発表が行われた。そして最終日はワークショップが行われた。また、毎晩その日の発表者によるポスターセッションが行われた。

今年の投稿数は 368 本中、最終的に受理されたのは 118 本 (採択率 32%) というところで、ほぼ例年通りであった。日本からは、筆者らを含めて 3 つのグループからの論文が発表された。全体としては、やはり統計的な考え方に基づく発表が多かった。個別の分野では、SVM をはじめとするカーネル法についての発表は全部で 35 本あり、引き続き盛り上がりつつある分野であるといえる。またそのうち 20 本はカーネル関数のデザインや、分類問題以外へのカーネル法の応用などに関する発表であり、カーネル法の研究の一層の広がりを感じた。近年数の多かった強化学習に関する発表は、若干減った印象ではあったが、行動から報酬を求めようという逆強化学習や、マルチェージェント系の強化学習のセッションが設けられるなど、新たな展開を感じさせた。また、応用のセッションとしては、招待講演の 2 件に対応して、バイオインフォマティクスやコンピュータビジョンのセッションが設けられた。

2. チュートリアル

Ghahramani による “Bayesian Methods for Machine Learning” では、通常、非常に大きな計算量になってしまうベイズ統計に基づく学習法に対し、サンプリングや変分法に基づき効率的に近似解を求める手法の最新動向が解説された。また、Gärtner による、“Kernels for Structured Data” では、配列や木、グラフで表現される構造データを、カーネル法によって扱うための種々のアプローチが解説された。各種手法を、データ間の関係を表現した「外的構造」と、データ内の要素間の関係を表現した「内的構造」に分類して紹介した点が非常に興味深かった。

その他、スペクトル手法によるクラスタリング、ゲーム理論に基づく学習や、eメールのフィルタリング手法、確率的論理などの解説が行われた。

3. 招待講演

招待講演では、機械学習以外の分野から、3 名の著名な研究者を迎え行われた

Myers のバイオインフォマティクスについての講演 “Whole Genome Sequencing, Comparative Genomics, and Systems Biology” は、彼が Celera Genomics 社で行った DNA のシーケンシングの話にはじまり、ゲノム比較による遺伝子発見法、そして最後に生物をシステムとして捉えようというシステムバイオロジーの話に発展した。また、ノーベル経済学賞受賞者である Engle による “Models of Dynamic Uncertainty in Univariate and Multivariate Systems” では、自身の ARCH モデルの研究を含めた経済時系列解析で用いられる種々のモデルの解説が行われた。特に興味深かったのはコンピュータビジョンの Black の講演 “Learning to See People” の、マーカーを使わないモーションキャプチャの話題であった。このタスクではビデオカメラで撮った映像から 30 以上の変数で指定される人間の姿勢を抽出しなければならないが、まず部分部分を別々に検出してから姿勢を推定しようとするという問題が多いという問題がある。そこで、四肢のつながりなどの人間の形についての制約を取り込んだグラフィカルモデルを定義し、全体をまとめて最適化し予測しようとすることで性能を改善しようという試みは非常に興味深かった。

4. 一般の発表

以下、一般の発表で筆者が面白いと感じた発表をいくつか紹介していくことにする。

まず、教師なし学習では Weinberg らの “Learning a Kernel Matrix for Nonlinear Dimensionality Reduction” が新しい非線形次元削減手法を提案した。変換前と変換後で、互いに近傍にあるデータ同士の距離が変わらないような制約のもとで、変換後の全点間距離の和が最大になるようにする (3 次元空間中で丸まった布を引っ張って広げるイメージ) ような非線形変換を求めることで、低次元の空間でデータを近似するという手法である。この問題は変換をカーネル行列によって暗に表現することによって半正定値計画法になり、そこそこの計算量で最適解を求めることが出来るというもので、実験も、壺を 360° 回転させながら撮った画像の集合が、見事に 2 次元



図 1 会議の開催された Banff Park Lodge

に落ちる様は印象的であった。この論文は distinguished paper を受賞していた。

同じく distinguished paper であった、Nguyen らの “Decentralized Detection and Classification using Kernel Method” では、分散したセンサーによってデータを収集し、情報を要約しながら最終的な分類予測を行うような学習器のカーネル法による実現法が提案された。今後、あらゆる場所にセンサーが設置され、膨大なデータが収集・解析されるような世の中になっていくとすると、それらの膨大なデータを一旦中心に集めてから解析を行うような非効率的なやり方でなく、この研究のように、できるだけ局所的にデータ処理を行い、うまく纏めながら解析していくような解析手法が有望であると思われる。

従来の分類学習の流れでは、Dekel らの “Large Margin Hierarchical Classification” において、Yahoo! などのサーチエンジンで行われているような Web ページの階層的なクラス分類の新しい手法が提案された。階層を無視した多クラスの分類と、クラス階層を上から分類していくやりかたの良いところを組み合わせた中間的な定式化を行い、このモデルのオンライン学習アルゴリズムを示した。

また、Ling らの “Decision Trees with Minimal Costs” では、誤分類コストだけでなく、属性を測るコストがあるような決定木が提案された。このようなモデルは、例えば医療分野などでの応用など様々な可能性があるのは勿論のこと、属性の計測コストが枝刈りの役割を果たし、過学習を防ぐよう働いているという点も非常に興味深い。

ところで、近年のひとつの大きな流れとして、入力データとして予めベクトル型で表現されておらず、配列や、木、グラフなどで表現された構造をもったデータを扱えるように学習手法を拡張する試みが精力的に行われている。さらに、最近は入力だけでなく出力も構造をもつようなケースが精力的に研究されている。こういった問題は自然言語処理やバイオインフォマティクスをはじめとして多くの分野で見られ、応用面でも非常に有望である。出力の構造が予め与えられている場合 (例えば入力と出力が同一の構造をもつ場合) には、構造に含まれる各変数にラベルを割り当てていく、構造ラベリングの問題と呼ばれ、自然言語処理における単語品詞付けや、バイオインフォマティクスにおけるタンパク質のアミノ酸配列の2次構造予測などがこれに当たる。また、より一般的には出力の構造も予測する問題が考えられ、例えば単語列を入力として木構造を出力する構文解析などがある。このような構造同士のマッピングを扱うための確率モデルとしては、2001年のICMLでLaffertyらによって提案された条件付き確率場 (Conditional Random Field) というモデルが広く受け入れられつつあるが、今回のICMLでは、条件付き確率場の拡張や効率的な学習法の提案などが多く見られた。

まず、大御所 Dietterich らの “Training Conditional Random Fields via Gradient Tree Boosting” では、あまり高速とはいえない条件付き確率場の学習を、回帰木によってモデルを近似することで gradient tree boosting という手法によって高速に学習する手法を提案した。

そして、Lafferty らの “Kernel Conditional Random Fields: Representation and Clique Selection” や Altun らの “Gaussian Process Classification for Segmenting and Annotating Sequences” では、カーネル化された条件付き確率場が提案された。現在この分野で注目を浴びているカーネル法と条件付き確率場が今回のICMLで融合、それも2グループによって同時に発表されたことは大変興味深い。



図 2 Banff の街並み

また SVM などで行われるマージン最大化に基づいて (カーネル化された) 条件付き確率場を学習する方法としては、昨年の ICML で、隠れマルコフ SVM という構造ラベリング学習法が提案されているが、今回は同グループ (Tsochantaridis ら) の “Support Vector Machine Learning for Interdependent and Structured Output Spaces” において、出力も構造を持つ場合にも適用可能にする拡張が行われた。

一方、以上で述べたカーネル化された条件付き確率場では、基本的には定数サイズの特徴しか考慮することができず、任意のサイズの特徴を効率的に取り込むようなカーネル法 (畳み込みカーネルなど) を用いることは従来不可能であった。この問題に対し筆者らによる “Kernel-Based Discriminative Learning Algorithms for Labeling Sequences, Trees and Graphs” では、固定サイズの特徴を用いる構造ラベリングのモデルと、任意サイズの特徴を扱う分類モデルをうまく結びつけることで解決を図っている。

5. おわりに

以上、簡単ではあったが ICML 2004 の概略を紹介した。なにぶん筆者の嗜好に基づいているため、教師あり学習に偏った紹介になってしまったことをお許しいただきたい。なお、ICML の Web サイト*1ではすべての論文が PDF でダウンロード可能となっているので、詳細や他の発表に興味を持たれた方は是非そちらを当たって頂きたい。

ICML は機械学習の分野では最も注目を浴びる、権威ある会議のひとつであるが、KDD などデータマイニング系の会議の full paper が軒並み超人的な倍率であるのに対し、ICML は約 3 倍と目標として高すぎず低すぎず、かつ十分なクオリティを保証するちょうど良い難しさであることに加え、内容的にも理論と応用のバランスのとれた、良質な会議であると思う。今後、この会議に日本からの発表がもっと増え、この分野での日本の研究者の認知度がより高まることを期待したい。

[鹿島 久嗣 (日本 IBM 東京基礎研究所)]

*1 <http://www.aicml.cs.ualberta.ca/~banff04/icml/>