COMS 6998-01 Advanced Machine Learning
Class Project
March 14, 2002
Prof. Tony Jebara

The class project will demonstrate your ability to use the material we explored in class and also go beyond it either via novel applications, theoretical development or new techniques. You can take on the project either individually or in teams of two. If you have a team of two people, the project should naturally be larger and more elaborate. The objective of the project is to work towards a conference-style paper (not necessarily one that would be publishable at this stage, but just an initial step).

Unlike the assignments, for the projects there is no fixed recipe to follow. Rather, you are free to pick a topic and direction that you find motivating. Your project can be in your current research area but make sure to show what novel work you have done for it for this course, you can't just submit previous work. Also, at the end of this page there is a list of papers and previous research efforts which you can consult as a starting point. Either consider implementing the algorithms in one of the papers below and using them on unusual datasets of your choice or consider extending the techniques in creative ways (i.e. note some of their limitations and show how to overcome them). Feel free to also bring new papers to the list below and suggest them as well. Places to look for papers include recent machine learning conferences such as NIPS [1], UAI [2], ICML [3], CVPR [4] and machine learning journals like JMLR [5], JAIR [6], Machine Learning, Pattern Recognition, Neural Computation, PAMI [7], and so forth. Many recent articles from these compilations are available online or in the library. You can find copies of the papers (postscript and pdf) through Citeseer, a popular search engine for computer science publications:

`http://citeseer.nj.nec.com/cs`

The following are important dates to remember:

April 1st: Email or hand in an abstract of your project: what you intend to do, theoretical or practical implications, relation to the class material, references to related publications and an idea of the results you expect, etc. This should be a one page (under 500 words) abstract and title with your name(s) and so forth. This will enable me to gauge the project, how ambitious it is and give you a little bit of guidance earlier on.

May 6th: In class, you will present the project briefly. Due to the class size and time limitations, your presentation should be limited to under 5 minutes (use about 5 slides). CVN students unfortunately cannot present so they should only send the slides which will be shown as posters for the rest of the class to browse (either as hardcopy or as a web page). Also, each person/team will hand in a conference-style paper describing the project in further detail. The paper should be on the order of 5 pages single spaced (either two-column or one-column). The work will be evaluated on its originality, its relevance, its technical soundness and its experimental thoroughness. Your paper should contain the following main ingredients or sections: problem definition, background work, your method, experiments, conclusions, future work ideas and references (at least 3). Again, refer to the listed papers, conference proceedings and journals for ideas on format and content.

[1] Neural Information Processing Systems
[2] Uncertainty in Artificial Intelligence
[3] International Conference on Machine Learning
[4] Computer Vision and Pattern Recognition
[5] Journal of Machine Learning Research
[6] Journal of Artificial Intelligence Research
[7] IEEE Transactions on Pattern Analysis and Machine Intelligence

The following are suggested topics and links to stimulate ideas if you don't already have a specific plan or area of interest:

Separating style and content with bilinear models
J. Tenenbaum and W. Freeman, http://www.merl.com/reports/docs/TR99-04.pdf

Variational bayes for mixture models
H. Attias, http://research.microsoft.com/~hagaia/uai99.ps
W. Penny, http://www.fil.ion.ucl.ac.uk/~wpenny/publications/vgbmm.ps

The relevance vector machine
M. Tipping, ftp.research.microsoft.com/users/mtipping/rvm_nips.ps.gz

Coupled hidden markov models for modeling interacting processes
M. Brand, http://www.media.mit.edu/people/brand/papers/brand-chmm.ps.gz

Estimating mixture models of images and inferring spatial transformations using EM
B. Frey and N. Jojic, http://www.psi.toronto.edu/~frey/papers/tmg-cvpr99.ps.Z

Solving multiclass learning problems via error-correcting output codes
T. Dietterich and T. Bakiri, ftp.cs.orst.edu/pub/tgd/papers/jair-ecoc.ps.gz

ICA: Edges are the 'independent components'of natural scenes
A. Bell and T. Sejnowski, ftp.cnl.salk.edu/pub/tony/edge.ps.Z

Transductive inference for text classification using SVMs
T. Joachims, http://www-ai.cs.uni-dortmund.de/DOKUMENTE/Joachims_99c.ps.gz

Feature selection for SVMs
J. Weston et al., http://www.ai.mit.edu/people/sayan/webPub/feature.ps

Kernel PCA and de-noising in feature spaces
S. Mika, et al., http://www.kernel-machines.org/papers/MikSchSmoMueRaeSch99.ps.gz

Regularization theory and neural networks architectures
F. Girosi and T. Poggio, http://www.ai.mit.edu/projects/cbcl/publications/ps/GirJonPogNC.ps.gz

Parameter estimation for linear dynamical systems
Z. Ghahramani, ftp.cs.toronto.edu/pub/zoubin/tr-96-2.ps.gz

GTM: The generative topographic mapping
C. Bishop, http://www.ncrg.aston.ac.uk/Papers/postscript/NCRG_96_015.ps.Z

Learning and classification of complex dynamics
B. North et al., http://www.ai.mit.edu/~jrennie/trg/papers/north-dynamics.pdf

Nonlinear image interpolation using manifold learning
C. Bregler and S. Omohundro, Advances in Neural Information Processing Systems 7, 1995

Scholkopf, et. al. Estimating the Support of a High-Dimensional
Distribution. Microsoft Technical Report, MSR-TR-99-87. 1999.

Nonlinear dimensionality reduction by locally linear embedding
S. Roweis and L. Saul, http://www.sciencemag.org/cgi/reprint/290/5500/2323.pdf

A generalization of principal component analysis to the exponential family
M. Collins et al, http://www.research.att.com/~dasgupta/pca.pdf

Mapping a manifold of perceptual observations
J. Tenenbaum, http://www-psych.stanford.edu/~jbt/man_nips.pdf

Learning internal representations
J. Baxter, http://www.syseng.anu.edu.au/~jon/papers/colt95.ps.gz

Probabilistic latent semantic analysis
T. Hofmann, http://www.cs.brown.edu/people/th/papers/Hofmann-UAI99.pdf

Model selection for support vector machines
O. Chapelle and V. Vapnik, http://www-connex.lip6.fr/~chapelle/ms_nips99.ps

Text classification from labeled and unlabeled documents using EM
K. Nigam et al, http://www-2.cs.cmu.edu/~mccallum/papers/emcat-mlj2000.ps.gz

Multivariate information bottleneck
N. Friedman et al, http://www.cs.huji.ac.il/~noamm/publications/UAI2001.ps.gz

Maximum Entropy Discrimination
T. Jaakkola et al, http://www.ai.mit.edu/~tommi/papers/maxent.ps

String matching kernels for text classification
H. Lodhi et al, http://www.support-vector.net/papers/string.ps

Generalized belief propagation,
J. Yedidia et al, http://www.ai.mit.edu/people/wtf/ijcai_final.pdf

A Bayesian approach to learning Bayesian networks with local structure
M Chickering et al, ftp.research.microsoft.com/pub/tr/tr-97-07.ps

Factorial hidden Markov models
Z. Ghahramani and M. Jordan, ftp://ftp.cs.toronto.edu/pub/zoubin/fhmmML.ps.gz

The EM algorithm for mixtures of factor analyzers
Z. Ghahramani and G. Hinton, ftp.cs.toronto.edu/pub/zoubin/tr-96-1.ps.gz

Gender classification with support vector machines
B. Moghaddam and M. Yang, http://www.merl.com/papers/docs/TR2000-01.pdf

Nonlinear prediction of chaotic time series using support vector machines
S. Mukherjee et al, http://www.ai.mit.edu/people/girosi/home-page/nnsp97.pdf

Alternatively, you could go off in more theoretical directions, i.e. devise new graphical models, variations

of mixture models, extensions and reapplications of the EM algorithm, or consider variations of classification, such as regression, partially labeled data, anomaly detection, etc. Applied efforts could go off into areas like image recognition, bioinformatics, text classification, speech, language modeling, reinforcement learning for games, audio classification, and so forth.

Potential datasets on which to try some of your learning algorithms:

`http://www1.ics.uci.edu/~mlearn/MLRepository.html`
`http://www.cs.toronto.edu/~delve/`
`http://www-personal.buseco.monash.edu.au/~hyndman/TSDL/`