

RAVE HARPAZ

EDUCATION

Ph.D. Computer Science, The Graduate Center, City University of New York 2008

- **Thesis Title:** Model-Based Linear Manifold Clustering
- **Thesis Adviser:** Distinguished Professor Robert M. Haralick
- **Research Interests:** Statistical pattern recognition, machine learning, data mining.
- **Applications:** Gene expression microarray analysis, text mining, financial modeling, machine vision.
- **GPA:** 3.9

LL.B. Law, Tel-Aviv University, School of Law, Tel Aviv, Israel 1997

- Appointed lawyer by the Israel Bar Association in 1998.

EXPERIENCE

Risk Quant Modeler, Merrill Lynch & Co., Inc., Basel II and Risk Analytics Group 2007 – present

- **Credit portfolio risk models-** Development, implementation, maintenance, and validation of “Jump to Default” (incremental) risk models of structured credit (CDS, CDO, FTD) and Mortgage CDO portfolios. Methods and concepts used: Merton model, single factor Gaussian copula, Vasicek loss distribution, Monte Carlo simulation, variance reduction, beta & extreme value distribution fitting, asset and default correlation.
- **Credit rating models** - Design, implementation, validation, and presentation of credit rating models based on default prediction and agency ratings. Segments worked on: small to medium companies, large corporations, mutual funds. Methods used: linear & logistic (constraint) regression, ROC analysis, hypothesis testing (Spearman rank correlation, Wilcoxon signed rank, Spiegelhalter), transition matrices, bootstrapping.

Research Engineer, Career360, LLC (Internet Startup) 2007

Lead developer of a search engine designed to retrieve and match job descriptions with resumes, which is based on Latent Semantic Analysis (concept matching using a statistical procedure) and a proprietary scoring system. Development and analysis was done using C++ and MATLAB, including approximately 4K lines of code.

Research Engineer, New York City Department of Finance's Office of Tax Policy 2004 - 2005

Designed and developed an interactive auditing system used by accountants to verify samples of corporate tax returns. The front-end of the system, Excel, interfaced a SAS tax return model. Implementation of the system rendered previous manual auditing procedures obsolete and increased team productivity and accuracy by a significant factor. System was implemented in VBA Excel and SAS.

Research Assistant, City University of New York 2003 – 2004

Grants:

- **Alphatech Inc.** – “Fast Relational Matching of 3D Objects”. Assisted my advisor with the research and development of a system used to identify E3D point clouds of military vehicles using the Relational Matching paradigm.
- **Raytheon** – “Estimating High Dimensional Probability Distributions and Target Recognition”. Assisted my advisor with the research and development of an algorithm to estimate probability distributions in large feature spaces. Based on the density estimates obtained by this algorithm, a Bayesian classifier was designed in order to classify target (military vehicles) and clutter pixels of gray-scale images.

Law Trainee, Tel-Aviv Court of Law, Tel-Aviv, Israel 1997 - 1998

Performed legal research in criminal law, assisted the Judge in the writing of sentences and verdicts.

TEACHING EXPERIENCE

Brooklyn College, Computer Science Department

2000 - 2006

Course instructor, teaching over 25 graduate and undergraduate classes in computer science.

- **Courses:** Data Structures Using C++, Discrete Mathematics, Advanced Programming Techniques Using C, Introduction to the Internet and Web-page Design, Assembly Programming.

AWARDS AND HONORS

- Mina Rees Dissertation Fellowship in the Sciences, award for the most promising Ph.D. dissertation, Graduate Center, City University of New York.
- Graduate Fellowship, Computer Science Department Graduate Center, City University of New York.
- Jack Wolfe Fellowship for superior achievement, Brooklyn College, City University of New York.

PROFESSIONAL ACTIVITIES

Reviewer for the journals "*IEEE Transactions on Knowledge and Data Engineering*", and "*Pattern Recognition Letters*".

PUBLICATIONS

- **Linear manifold clustering in high dimensional spaces by stochastic search** (with Robert Haralick), *Pattern Recognition* (2007), vol. 40(10), pp 2672-2684.
- **Linear Manifold Correlation Clustering** (with Robert Haralick), Invited Paper, *International Journal of Information Technology and Intelligent Computing* (2007), vol 2, no. 2 .
- **Model-based Subspace Correlation Clustering** (with Robert Haralick), *Pattern Recognition* (2008), accepted pending minor revision.
- **Modeling High-Dimensional Probability Distributions via Linear Manifold Clusters** (with Robert Haralick), *Pattern Recognition Letters* (2008), pending revision.
- **Mining Subspace Correlations** (with Robert Haralick), *In Proceedings of the IEEE Symposium on Computational Intelligence and Data Mining (CIDM 2007)*, pp 335-342.
- **Exploiting the Geometry of Gene Expression Patterns for Unsupervised Learning** (with Robert Haralick), *In Proceedings of the 18th International Conference on Pattern Recognition (ICPR 2006)*, vol. 2 pp 670-674.
- **Linear Manifold Clustering** (with Robert Haralick), *In Proceedings of the International Conference on Data Mining and Machine Learning (MLDM 2005)*, Lecture Notes in Computer Science, Springer Verlag LNAI 3587 pp 132-141.
- **Linear Manifold Embedding of Pattern Clusters** (with Robert Haralick), *DIMACS Workshop on Detecting and Processing Regularities in High Throughput Biological Data*, 2005.
- **The EM Algorithm as a Lower Bound Optimization Technique**, *Technical Report TR-2006001*, Graduate Center, City University of New York, 2006.
- **Independent Component Analysis: An Introduction**, *Technical Report TR-2007007*, Graduate Center, City University of New York, 2005.
- **Fast Relational Matching**, *Technical Report TR-1401*, ALPHATECH Inc., 2003.
- **The Satisfiability Problem- From The Theory of NP-completeness to State-Of-The-Art SAT Solvers**, *Technical Report TR-2007008*, Graduate Center, City University of New York, 2003.

PROGRAMMING LANGUAGES

- Extensive C/C++, MATLAB, VBA, some experience with SAS and SQL.
- Scientific Libraries: Numerical Recipes in C, GNU Scientific Library, GNU Multiple Precision Arithmetic, SVDPACK.