SAMY BENGIO

Work address:
Apple
One Apple Park Way
Cupertino, CA 95014, USA
bengio@apple.com
http://bengio.abracadoudou.com/

Languages: french and english
Citizenship: canadian

Education

Doctorate: Ph.D. in Computer Science — Université de Montréal, 1989–1993

Subject: *Optimization of a Parametric Learning Rule for Neural Networks.*

Master: M.Sc. in Computer Science — Université de Montréal, 1986–1989

Subject: *Integration of Traditional and Intelligent Tutoring Systems.*

Diploma: B.Sc. in Computer Science — Université de Montréal, 1983–1986

Professional Experience

Apple Inc — 2021–

Description: Consumer electronics, computer software and online services.
Position: Senior Director, AI and Machine Learning Research.
Activity: Fundamental research on machine learning algorithms.
Main Research Tracks Long term curiosity-based foundational research in machine learning.

Google Inc — 2007–2021

Description: Search Engine and Advertising Company.
Position: Research Scientist in Machine Learning.
Activity: Fundamental and applied research on machine learning algorithms.
Main Research Tracks Deep Learning; Ranking algorithms for large scale problems; Applications to images, videos, speech, text, etc.

IDIAP Research Institute — 1999–2007

Description: Swiss research center affiliated to EPFL devoted to multimodal interactions, speech processing, vision processing, and machine learning.
Position: Senior Researcher, Head of the Machine Learning group.
Activity: Fundamental and applied research on machine learning algorithms. Supervision of Ph.D. students and postdoctoral fellows. European and Swiss national research projet proposal and management. Organization of international workshops and conferences. Teaching at the doctoral level. Development of machine learning tools. Collaboration with other IDIAP and EPFL researchers (in speech, vision and multimodal processing).
Main Research Tracks Various complex graphical models for speech and multimodal/multi-channel analysis; Various text representation and information retrieval approaches; Efficient algorithms for large scale databases; Various person authentication approaches.

Microcell Labs 1997–1999

Description: Research center of a Canadian mobile telecommunication company using the GSM standard.

Position: Research Director.

Activity: Direction of a research team focusing on the development and the application of various learning algorithms to exploit many dataware houses regarding the internal processes and client behaviors of a telecommunication company. Problems tackled: fraud dectection, churning and traffic prediction.

Centre Interuniversitaire de Recherche en ANalyse des Organisations 1996–1997,1999

Description: Canadian academic research center with many industrial partnerships devoted to the economical aspects of the organizations.

Position: Researcher.

Activity: Development of various machine learning and statistical models that make predictions and take decisions on financial markets: portfolio management, asset prediction, portfolio selection. In an other project, development of a machine learning model that simulates an agent on various types of auctions (buy at the best price given known information).

Institut National de la Recherche Scientifique - Télécommunications 1995–1996

Description: Canadian academic research center on telecommunications.

Position: Postdoctoral Fellow.

Activity: Development and use of an Asynchronous Input/Output Hidden Markov Model (which can also be viewed as a connectionist recurrent mixture of experts) for speech recognition. This model was particularly well suited for non-stationary temporal problems.

Centre National d'Études des Télécommunications, France Télécom 1994–1995

Description: Research center of the french national telecommunication company.

Position: Postdoctoral Fellow.

Activity: Time series prediction with Neural Networks. Series were taken from a physical domain (sunspot-related) as well as a telecommunication domain. Various mixture models were proposed.

Centre de Recherche sur les Transports, Université de Montréal 1993

Description: Canadian academic research center on all aspects of transports.

Position: Research Assistant.

Activity: Use of Genetic Algorithms and Tabu Search for a Combinatorial Optimization Problem: *the Vehicle Routing Problem with Time Windows*.

Computer Science Technical Knowledge

- **Development of several simulators** of machine learning algorithms, including many types of neural networks, support vector machines, [input/output] hidden Markov models, Gaussian mixture models, and other statistical models, genetic algorithms, genetic programming, simulated annealing, tabu search, and expert systems. See for instance `http://www.Torch.ch`.
- Many years experience in system administration on multi-platform environments (Linux, Macintosh, Unix, Windows, etc).
- **Known computer languages:** C, C++, Java, Lisp, Mathematica, Matlab, Objective C, Pascal, Prolog, Splus, Smalltalk.

Doctoral Students

PhD thesis obtained under joint supervision:

- Françoise Fessant, Université de Rennes, 1995
- Miguel Moreira, EPFL Lausanne, 2001
- Nicolas Gilardi, Université de Lausanne, 2002
- Katrin Weber, EPFL Lausanne, 2003
- Ronan Collobert, Université de Paris, 2004
- Fabien Cardinaux, EPFL Lausanne, 2005
- Silvia Chiappa, EPFL Lausanne, 2006
- Johnny Mariéthoz, Université de Lyon 2, 2006
- Norman Poh, EPFL Lausanne, 2006
- Alexei Pozdnoukhov, EPFL Lausanne, 2006
- Mikaela Keller, EPFL Lausanne, 2006
- Christos Dimitrakakis, EPFL Lausanne, 2006
- Yann Rodriguez, EPFL Lausanne, 2006
- Dong Zhang, EPFL Lausanne, 2006
- David Grangier, EPFL Lausanne, 2008
- Jean-Francois Paiement, EPFL Lausanne, 2008

Member of other PhD thesis juries:

- Sébastien Marcel, Université de Rennes, 2000
- Pierre-Edouard Sottas, EPFL Lausanne, 2002
- Todd Stephenson, EPFL Lausanne, 2003
- Torsten Butz, EPFL Lausanne, 2003
- Liva Ralaivola, Université de Paris 6, 2003
- Vlad Popovici, EPFL Lausanne, 2004
- Mathieu Ben, Université de Rennes, 2004
- Serghei Kosinov, Université de Genève, 2005
- Jean-Julien Aucouturier, Université de Paris 6, 2006
- Sylvain Ferrandiz, Université de Caen, 2006
- Gianluca Monaci, EPFL Lausanne, 2007
- Ivana Arsic, EPFL Lausanne, 2007
- Marie Szafranski, Université de Technologie de Compiègne, 2008
- Mihai Gurban, EPFL Lausanne, 2008
- Jie Luo, EPFL Lausanne, 2011

- Majid Yazdani, EPFL Lausanne, 2013
- Pierre-Michel Bousquet, Université d'Avignon, 2014
- Seungyeon Kim, Georgia Tech, 2015
- Dimitri Palaz, EPFL Lausanne, 2016
- Hieu Pham, CMU, 2021
- Paul Best, University of Toulon, 2022

Member of other academic juries:

- Hervé Glotin, HDR, Université Sud Toulon Var, 2007
- Vincent Lemaire, HDR, Université de Paris Sud, 2008

Teaching Activities

EPFL Lausanne 2005-2006

Activity: Course **IC-49**, on *Statistical Machine Learning from Data* given in the *I&C Computer, Communication and Information Sciences Doctoral Program*

- Lecture notes available at `http://bengio.abracadoudou.com/lectures`.

IDIAP Research Institute 2002-2003

Activity: Advanced lectures on *statistical machine learning*

- Lecture notes available at `http://bengio.abracadoudou.com/lectures/old`.

Université de Montréal 1992,1993,1996

Activity: Teaching replacement for M.Sc./Ph.D. courses.

- Artificial Neural Networks.
- Genetic Algorithms.
- Hidden Markov Models.

Activity: Spring School (1996) on Learning Algorithms.

- Introduction to (Input/Output) Hidden Markov Models.

Funding Record

Swiss National Science Foundation Projects:

IM2.{ACP,MI,BMI,MPR} – Interactive Multimodal Information Management.
4 PhD and 2 postdoc, 280kCHF/year for 10 years.

MULTI – Multimodal Interaction and Multimedia Data Mining.
several PhD thesis, 290kCHF/year for 3 years.

KerSpeech – Kernel Methods for Speech and Video Sequence Analysis.
1 PhD, 2 years, in the context of the Indo Swiss Joint Research Programme (ISJRP).

ADASEQ – Ensemble Methods for Sequence Processing.
1 PhD, 40kCHF/year for 4 years.

CARTANN – Cartography by Artificial Neural Networks.
1 PhD thesis finished, 40kCHF/year for 4 years.

Divide and Learn I – Various mixture models proposed.
1 PhD thesis finished, 40kCHF/year for 4 years.

Divide and Learn II – Mixture models for large datasets.
2 PhD, 100kCHF/year for 4 years done.

GLAD – Use of Boolean Methods for Classification.
1 PhD finished, 40kCHF/year for 4 years.

KERNEL – Kernel Methods for Sequence Processing.
1 PhD, 40kCHF/year for 4 years.

SCRIPT – Cursive Handwriting Recognition.
1 PhD thesis finished, 40kCHF/year for 4 years.

European Projects

PASCAL – Pattern Analysis, Statistical Modelling and Computational Learning, 6th Framework Programme, Information Society Technology, Network of Excellence, 6th Framework Programme, Information Society Technology, Network of Excellence, 4 years. Member of the steering committee.

BANCA – Biometric Access Control for Networked and e-Commerce Applications, 5th Framework Programme, Information Society Technology.
2 researchers, 200kCHF/year for 3 years.

EDAM – Environmental data mining: machine Learning algorithms and statistical tools for monitoring and forecasting, INTAS foundation.
1 invited researcher, 40kCHF/year for 2 years.

LAVA – Learning for Adaptable Visual Assistants.
1 postdoc and 2 PhD, 200kCHF/year for 3 years.

COST-275 – Biometric-Based Recognition of People over the Internet.
1 PhD, 40kCHF/year for 3 years.

Innovation, Technology Transfer, Patents

Patents

- S. Bengio, F. Clerot, A. Gravey, D. Collobert. Process for the renewal -during communication- of the traffic parameters of an ATM network. France Telecom. United States Patent 6018517. International Patent WO 97/31456. 1997.

Technology Transfer

- Release of the popular Support Vector Machine software SVMTorch on the web for academic research (see `http://bengio.abracadoudou.com/SVMTorch.html`).
- Release of the complete machine learning library Torch on the web (see `http://www.Torch.ch`).

Other Professional Activities

Action Editor for the Journal of Machine Learning Research since 2012

Member of the Editorial Board for the Journal of Machine Learning Research, 2009-2012

Member of the Editorial Board for the Machine Learning journal, 2010–2020

Associate Editor for the Journal of Computational Statistics, 2002-2011

Associate Editor for the Journal of Selected Topics in Signal Processing, 2009

Member of the Scientific Advisory Board of the Simons Institute for the Theory of Computing, 2018–2020

Member of the External Advisory Board of the NSF ID4: Institute for Data Driven Dynamical Design, 2022–

Member of the University of Montreal International Advisory Committee to the Rector, 2022–

Organization of Conferences and Workshops:

- Member of the Board:
  - ICLR (2018-2020)
  - NeurIPS (2019-)
- Chair:
  - NeurIPS: Ethics review chair (2021).
  - NeurIPS: General chair (2018); senior programme chair (2017).
  - ICLR: co-programme chair (2015, 2016).
  - BayLearn: a new Workshop in Machine Learning in the Bay Area (BayLearn'2012-2016).
  - BigVision 2014, 2015: a CVPR Workshop on Big Data for Computer Vision (CVPR'2014, 2015).
  - BigVision 2012: a NIPS Workshop on Big Data for Computer Vision (NIPS'2012).
  - NIPS Demonstration Chair (NIPS'2011)
  - NIPS Workshop on Efficient Machine Learning (NIPS'2007)
  - NIPS Workshop on Learning to Compare Examples (NIPS'2006)
  - NIPS Workshop on Multimodal Signal Processing (NIPS'2004)
  - WebVision: ECCV Workshop on Computer Vision for the Web (ECCV'2012)
  - Workshop on Multimodal Interaction and Related Machine Learning Algorithms (MLMI'2004, 2005, 2006)
- Programme Committee Chair - Senior Area Chair:
  - International Conference on Learning Representations (ICLR'2022, '2023, '2024)
  - International Conference on Machine Learning (ICML'2019, '2021, '2022, '2023)
  - Neural Information Processing Systems (NeurIPS'2020, '2022, '2023)
- Programme Committee Chair - Area Chair:
  - European Conference on Machine Learning (ECML-PKDD'2010)
  - IEEE Workshop on Neural Networks for Signal Processing (NNSP'2002)
  - International Conference on Learning Representations (ICLR'2019, '2021)
  - International Conference on Machine Learning (ICML'2009, 2012, 2015, 2016, 2020)
  - International Joint Conference on Artificial Intelligence (IJCAI'2011), Senior Program Committee
  - International Joint Conference on Artificial Intelligence (IJCAI'2005), Poster Track
  - Neural Information Processing Systems (NIPS'2003, 2006, 2012, 2014, 2015)
- Reviewer - Programme Committee Member - International Conferences:
  - European Symposium on Artificial Neural Networks (ESANN)
  - Extraction et Gestion des Connaissances (EGC)
  - IEEE Conference on Face and Gesture Recognition (FG)
  - IEEE Conference on Machine Learning for Signal Processing, (NNSP, MLSP)
  - IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)
  - IEEE International Conference on Robotics and Automation (ICRA)
  - IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
  - International Conference on Artificial Intelligence and Statistics (AISTATS)
  - International Conference on Audio and Video Based Biometric Person Authentication (AVBPA)
  - International Conference on Biometrics (ICB)
  - International Conference on Computer Vision and Pattern Recognition (CVPR)
  - International Conference on Machine Learning (ICML)
  - International Conference on Pattern Recognition (ICPR)
  - International Conference on Biometrics Authentication (ICBA)
  - International Conference on Learning Representations (ICLR)
  - International Joint Conference on Neural Networks (IJCNN)

  - International Workshop on Multiple Classifier Systems (MCS)
  - Neural Information Processing Systems (NIPS, NeurIPS)
- Reviewer - Programme Committee Member - International Workshops:
  - AAAI Spring Symposium on Knowledge Representation and Reasoning (2015)
  - AAAI Spring Symposium on Human Behavior Modeling (2009)
  - ACM Symposium on Applied Computing - Special Track on Applied Biometrics (2010)
  - CVPR Workshop on Multi-biometrics (CVPR'2006)
  - CVPR Workshop on Beyond Patches (CVPR'2006)
  - International Workshop on Biometric Recognition Systems (IWBRS'2005)
  - ICML Extreme Classification Workshop (2015)
  - Morpho Challenge Workshop (2005, 2007)
  - SIGIR 2007 Workshop on Searching Spontaneous Conversational Speech
  - Workshop on Multimodal Interaction and Related Machine Learning Algorithms (MLMI'2007, 2008, 2009)
  - Multimodal User Authentication Workshop (MMUA'2006)
  - NIPS Workshop on Machine Learning for Implicit Feedback and User Modeling (NIPS'2005)
  - Spatial Interpolation Comparison (SIC'2004)
- Member of Advisory Boards: MLMI
- Session Chair: Eurospeech'2003, NNSP'2002, ICASSP'2006, ICML'2012

Reviewer for many Journals:

- Computer Speech and Language
- EURASIP Journal of Applied Signal Processing
- IEEE Transactions on Biomedical Engineering
- IEEE Transactions on Image Processing
- IEEE Transactions on Evolutionary Computation
- IEEE Transactions on Neural Networks
- IEEE Transactions on Pattern Analysis and Machine Intelligence
- IEEE Transactions on Speech and Audio Processing
- IEEE Transactions on Systems, Man and Cybernetics - Part B
- Information Fusion
- International Journal of Neural Systems
- International Journal of Pattern Recognition and Artificial Intelligence
- Journal of Machine Learning Research
- Journal of Pattern Recognition Letters
- Machine Learning Journal
- Neural Computing Surveys
- Neural Computation
- Neural Networks
- Neurocomputing Journal
- Pattern Recognition
- Pattern Recognition Letters
- Signal Processing
- Speech Communication

Publication List

Articles in Peer Reviewed Journals

[1] C. Zhang, S. Bengio, and Y. Singer. Are all layers created equal? *Journal of Machine Learning Research, JMLR*, 23:1–28, 2022.

[2] B. Kim, E. Reif, M. Wattenberg, S. Bengio, and M. C. Mozer. Neural networks trained on natural scenes exhibit gestalt closure. *Computational Brain and Behavior*, 4(3):251–263, 2021.

[3] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals. Understanding deep learning (still) requires rethinking generalization. *Communications of the ACM*, 64(3):107–115, 2021.

[4] S. Bengio, K. Dembczynski, T. Joachims, M. Kloft, and M. Varma. Extreme Classification (Dagstuhl Seminar 18291). *Dagstuhl Reports*, 8(7):62–80, 2019.

[5] J. Chorowski, R. J. Weiss, S. Bengio, and A. van den Oord. Unsupervised speech representation learning using wavenet autoencoders. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 27:2041–2053, 2019.

[6] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan. Show and tell: Lessons learned from the 2015 mscoco image captioning challenge. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI*, 39(4):652–663, 2017.

[7] J. Lee, S. Kim, G. Lebanon, Y. Singer, and S. Bengio. LLORMA: Local low-rank matrix approximation. *Journal of Machine Learning Research, JMLR*, 17:1–24, 2016.

[8] M. R. Gupta, S. Bengio, and J. Weston. Training highly multiclass classifiers. *Journal of Machine Learning Research, JMLR*, 15:1461–1492, 2014.

[9] S. Bengio, L. Deng, H. Larochelle, H. Lee, and R. Salakhutdinov. Guest editors' introduction: Special section on learning deep architectures. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI*, 35:1795–1797, 2013.

[10] H. Elmlund, D. Elmlund, and S. Bengio. PRIME: Probabilistic initial 3d model generation for single-particle cryo-electron microscopy. *Structure*, 21:1299–1306, 2013.

[11] C. Dimitrakakis and S. Bengio. Phoneme and sentence-level ensembles for speech recognition. *EURASIP Journal on Audio, Speech, and Music Processing*, 2011, 2011.

[12] J. Weston, S. Bengio, and P. Hamel. Multi-tasking with joint semantic spaces for large-scale music annotation and retrieval. *Journal of New Music Research*, 40:337–348, 2011.

[13] G. Chechik, V. Sharma, U. Shalit, and S. Bengio. Large scale online learning of image similarity through ranking. *Journal of Machine Learning Research, JMLR*, 11:1109–1135, 2010.

[14] D. Erhan, Y. Bengio, A. Courville, P.-A. Manzagol, P. Vincent, and S. Bengio. Why does unsupervised pre-training help deep learning? *Journal of Machine Learning Research, JMLR*, 11:625–660, 2010.

[15] R. F. Lyon, M. Rehn, S. Bengio, T. C. Walters, and G. Chechik. Sound retrieval and ranking using sparse auditory representations. *Neural Computation*, 22(9):2390–2416, 2010.

[16] J. Weston, S. Bengio, and N. Usunier. Large scale image annotation: Learning to rank with joint word-image embeddings. *Machine Learning Journal*, 81(1):21–35, 2010.

[17] J. Keshet, D. Grangier, and S. Bengio. Discriminative keyword spotting. *Speech Communication*, 51:317–329, 2009.

[18] J.-F. Paiement, S. Bengio, and D. Eck. Probabilistic models for melodic prediction. *Artificial Intelligence Journal*, 173(14):1266–1274, 2009.

[19] J.-F. Paiement, Y. Grandvalet, and S. Bengio. Predictive models for music. *Connection Science*, 21(2 & 3):253–272, 2009.

[20] D. Grangier and S. Bengio. A discriminative kernel-based model to rank images from text queries. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI*, 30(8):1371–1384, 2008.

[21] H. Paugam-Moisy, R. Martinez, and S. Bengio. Delay learning and polychronization for reservoir computing. *Neurocomputing*, 71(7–9):1143–1158, 2008.

[22] J. MariÃ©thoz and S. Bengio. A kernel trick for sequences applied to text-independent speaker verification systems. *Pattern Recognition*, 40:2315–2324, 2007.

[23] N. Poh, A. Martin, and S. Bengio. Performance generalization in biometric authentication using joint user-specific and sample bootstraps. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI*, 29(3):492–498, 2007.

[24] S. Sonnenburg, M. L. Braun, C. Soon Ong, S Bengio, L. Bottou, G. Holmes, Y. LeCun, K.-R. MÃ¼ller, F. Pereira, C. E. Rasmussen, G. RÃ¤tsch, B. SchÃ¶lkopf, A. Smola, P. Vincent, J. Weston, and R. Williamson. The need for open source software in machine learning. *Journal of Machine Learning Research, JMLR*, 8:2443–2466, 2007.

[25] F. Cardinaux, C. Sanderson, and S. Bengio. User authentication via adapted statistical models of face images. *IEEE Transactions on Signal Processing*, 54(1):361–373, 2006.

[26] N. Poh and S. Bengio. Database, protocol and tools for evaluating score-level fusion algorithms in biometric authentication. *Pattern Recognition*, 39(2):223–233, 2006.

[27] A. Pozdnoukhov and S. Bengio. Invariances in kernel methods: From samples to objects. *Pattern Recognition Letters*, 27(10):1087–1097, 2006.

[28] Y. Rodriguez, F. Cardinaux, S. Bengio, and J. MariÃ©thoz. Measuring the performance of face localization systems. *Image and Vision Computing*, 24(8):882–893, 2006.

[29] C. Sanderson, S. Bengio, and Y. Gao. On transforming statistical models for non-frontal face verification. *Pattern Recognition*, 39(2):288–302, 2006.

[30] D. Zhang, D. Gatica-Perez, S. Bengio, and I. McCowan. Modeling individual and group actions in meetings with layered HMMs. *IEEE Transactions on Multimedia*, 8(3):509–520, 2006.

[31] C. Dimitrakakis and S. Bengio. Online adaptive policies for ensemble classifiers. *Neurocomputing*, 64:211–221, 2005.

[32] J. MariÃ©thoz and S. Bengio. A unified framework for score normalization techniques applied to text independent speaker verification. *IEEE Signal Processing Letters*, 12(7):532–535, 2005.

[33] I. McCowan, D. Gatica-Perez, S. Bengio, G. Lathoud, M. Barnard, and D. Zhang. Automatic analysis of multimodal group actions in meetings. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI*, 27(3):305–317, 2005.

[34] N. Poh and S. Bengio. How do correlation and variance of base classifiers affect fusion in biometric authentication tasks? *IEEE Transactions on Signal Processing*, 53(11):4384–4396, 2005.

[35] V. Popovici, S. Bengio, and J.-P. Thiran. Kernel matching pursuit for large datasets. *Pattern Recognition*, 38(12):2385–2390, 2005.

[36] S. Bengio. Multimodal speech processing using asynchronous hidden markov models. *Information Fusion*, 5(2):81–89, 2004.

[37] F. de Wet, K. Weber, L. Boves, B. Cranen, S. Bengio, and H. Bourlard. Evaluation of formant-like features for automatic speech recognition. *Journal of the Acoustical Society of America, JASA*, 116(3):1781–1792, 2004.

[38] A. Vinciarelli, S. Bengio, and H. Bunke. Offline recognition of unconstrained handwritten texts using HMMs and statistical language models. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI*, 26(6):709–720, 2004.

[39] R. Collobert, Y. Bengio, and S. Bengio. Scaling large learning problems with hard parallel mixtures. *International Journal on Pattern Recognition and Artificial Intelligence, IJPRAI*, 17(3):349–365, 2003.

[40] K. Weber, S. Ikbal, S. Bengio, and H. Bourlard. Robust speech recognition and feature extraction using HMM2. *Computer, Speech and Language*, 17(2-3):195–211, 2003.

[41] S. Bengio, C. Marcel, S. Marcel, and J. MariÃ©thoz. Confidence measures for multimodal identity verification. *Information Fusion*, 3(4):267–276, 2002.

[42] R. Collobert, S. Bengio, and Y. Bengio. A parallel mixture of SVMs for very large scale problems. *Neural Computation*, 14(5):1105–1114, 2002.

[43] A. Vinciarelli and S. Bengio. Writer adaptation techniques in HMM based off-line cursive script recognition. *Pattern Recognition Letters*, 23(8):905–916, 2002.

[44] R. Collobert and S. Bengio. SVMTorch: Support vector machines for large-scale regression problems. *Journal of Machine Learning Research, JMLR*, 1:143–160, 2001.

[45] S. Bengio and Y. Bengio. Taking on the curse of dimensionality in joint distributions using neural networks. *IEEE Transaction on Neural Networks, special issue on data mining and knowledge discovery*, 11(3):550–557, 2000.

[46] N. Gilardi and S. Bengio. Local machine learning models for spatial data analysis. *Journal of Geographic Information and Decision Analysis*, 4(1):11–28, 2000.

[47] S. Bengio, Y. Bengio, J. Robert, and G. BÃ©langer. Stochastic learning of strategic equilibria for auctions. *Neural Computation*, 11(5):1199–1209, 1999.

[48] S. Bengio, F. Fessant, and D. Collobert. Use of modular architectures for time series prediction. *Neural Processing Letters*, 3(2):101–106, 1996.

[49] F. Fessant, S. Bengio, and D. Collobert. On the prediction of solar activity using different neural network models. *Annales Geophysicae*, 14:20–26, 1996.

[50] J.-Y. Potvin and S. Bengio. The vehicle routing problem with time windows - part II: Genetic search. *INFORMS Journal on Computing*, 8(2):165–172, 1996.

[51] S. Bengio, Y. Bengio, and J. Cloutier. On the search for new learning rules for ANNs. *Neural Processing Letters*, 2(4):26–30, 1995.

[52] S. Bengio, G. Brassard, Y. Desmedt, C. Goutier, and J.-J. Quisquater. Secure implementation of identification systems. *Journal of Cryptology*, 4(3):175–183, 1991.

Articles in Peer Reviewed Conferences

[53] E. Abbe, S. Bengio, A. Lotfi, C. Sandon, and O. Saremi. How far can transformers reason? the locality barrier and inductive scratchpad. *ArXiv*, 2406.06467, 2024.

[54] E. Boix-Adsera, O. Saremi, E. Abbe, S. Bengio, E. Littwin, and J. Susskind. When can transformers reason with abstract symbols. In *International Conference on Learning Representations, ICLR*, 2024.

[55] H. Zhou, A. Bradley, E. Littwin, N. Razin, O. Saremi, J. Susskind, S. Bengio, and P. Nakkiran. What algorithms can transformers learn: A study in length generalization. In *International Conference on Learning Representations, ICLR*, 2024.

[56] E. Abbe, S. Bengio, A. Lotfi, and K. Rizk. Generalization on the unseen, logic reasoning and degree curriculum. In *International Conference on Machine Learning, ICML*, 2023. Outstanding Paper Award.

[57] S. Abnar, O. Saremi, L. Dinh, S. Wison, M. A. Bautista, C. Huang, V. Thilak, E. Littwin, J. Gu, J. Susskind, and S. Bengio. Adaptivity and modularity for efficient generalization over task complexity. *ArXiv*, 2310.08866, 2023.

[58] D. Berrebbi, R. Collobert, S. Bengio, N. Jaitly, and T. Likhomanenko. Continuous pseudo-labeling from the start. In *International Conference on Learning Representations, ICLR*, 2023.

[59] E. Boix-Adsera, E. Littwin, E. Abbe, S. Bengio, and J. Susskind. Transformers learn through gradual rank increase. In *Advances In Neural Information Processing Systems, NeurIPS*, 2023.

[60] S. d'Ascoli, S. Bengio, J. Susskind, and E. Abbe. Boolformer: Symbolic regression of logic functions with transformers. *ArXiv*, 2309.12207, 2023.

[61] E. Abbe, S. Bengio, E. Cornacchia, J. Kleinberg, A. Lotfi, M. Raghu, and C. Zhang. Learning to reason with neural networks: Generalization, unseen data and boolean measures. In *Advances In Neural Information Processing Systems, NeurIPS*, 2022.

[62] T. Likhomanenko, R. Collobert, N. Jaitly, and S. Bengio. Continuous pseudo-labeling in ASR. *ArXiv*, 2211.06007, 2022.

[63] M. L. Iuzzolino, M. C. Mozer, and S. Bengio. Improving anytime prediction with parallel cascaded networks and a temporal-difference loss. In *Advances In Neural Information Processing Systems, NeurIPS*, 2021.

[64] Y. Jiang, P. Natekar, M. Sharma, S. K. Aithal, D. Kashyap, N. Subramanyam, C. Lassance, D. M. Roy, G. K. Dziugaite, S. Gunasekar, I. Guyon, P. Foret, S. Yak, H. Mobahi, B. Neyshabur, and S. Bengio. Methods and analysis of the first competition in predicting generalization of deep learning. In *Proceedings of Machine Learning Research*, volume 133, pages 170–190, 2021.

[65] Y. Li, S. Si, G. Li, C.-J. Hsieh, and S. Bengio. Learnable fourier features for multi-dimensional spatial positional encoding. In *Advances In Neural Information Processing Systems, NeurIPS*, 2021.

[66] C. Zhang, M. Raghu, J. Kleinberg, and S. Bengio. Pointer value retrieval: A new benchmark for understanding the limits of neural network generalization. *ArXiv*, 2107.12580, 2021.

[67] Y. Guo, J. Choi, M. Moczulski, S. Feng, S. Bengio, M. Norouzi, and H. Lee. Memory based trajectory-conditioned policies for learning from sparse rewards. In *Advances In Neural Information Processing Systems, NeurIPS*, 2020.

[68] S. Hooker, N. Moorosi, G. Clark, S. Bengio, and E. Denton. Characterising bias in compressed models. *ArXiv*, 2010.03058, 2020.

[69] Y. Jiang, P. Foret, S. Yak, D. M. Roy, H. Mobahi, G. K. Dziugaite, S. Bengio, S. Gunasekar, I. Guyon, and B. Neyshabur. NeurIPS 2020 competition: Predicting generalization in deep learning. *ArXiv*, 2012.07976, 2020.

[70] Y. Jiang, B. Neyshabur, H. Mobahi, D. Krishnan, and S. Bengio. Fantastic generalization measures and where to find them. In *International Conference on Learning Representations, ICLR*, 2020.

[71] Y. Li, J. Amelot, X. Zhou, S. Bengio, and S. Si. Auto completion of user interface layout design using transformer-based tree decoders. *ArXiv*, 2001.05308, 2020.

[72] C. Luo, H. Mobahi, and S. Bengio. Data augmentation via structured adversarial perturbations. *ArXiv*, 2011.03010, 2020.

[73] A. Raghu, M. Raghu, S. Bengio, and O. Vinyals. Rapid learning or feature reuse? towards understanding the effectiveness of MAML. In *International Conference on Learning Representations, ICLR*, 2020.

[74] C. Zhang, S. Bengio, M. Hardt, M. C. Mozer, and Y. Singer. Identity crisis: Memorization and generalization under extreme overparameterization. In *International Conference on Learning Representations, ICLR*, 2020.

[75] V. Birodkar, H. Mobahi, and S. Bengio. Semantic redundancies in image-classification datasets: The 10% you don't need. *ArXiv*, 1901.11409, 2019.

[76] V. Birodkar, H. Mobahi, D. Krishnan, and S. Bengio. A closed-form learned pooling for deep classification networks. *ArXiv*, 1906.03808, 2019.

[77] Z. Chen, Y. Li, S. Bengio, and S. Si. You look twice: Gaternet for dynamic filter selection in CNNs. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2019.

[78] W.-L. Chiang, X. Liu, S. Si, Y. Li, S. Bengio, and C.-J. Hsieh. Cluster-GCN: An efficient algorithm for training deep and large graph convolutional networks. In *Conference on Knowledge Discovery and Data Mining, KDD*, 2019. Frontiers of Science Award from ICBS 2023.

[79] D. Duckworth, A. Neelakantan, B. Goodrich, L. Kaiser, and S. Bengio. Parallel scheduled sampling. *ArXiv*, 1906.04331, 2019.

[80] Y. Guo, J. Choi, M. Moczulski, S. Bengio, M. Norouzi, and H. Lee. Efficient exploration with self-imitation learning via trajectory-conditioned policy. *ArXiv*, 1907.10247, 2019.

[81] Y. Jiang, D. Krishnan, H. Mobahi, and S. Bengio. Predicting the generalization gap in deep networks with margin distributions. In *International Conference on Learning Representations, ICLR*, 2019.

[82] B. Kim, E. Reif, M. Wattenberg, and S. Bengio. Do neural networks show gestalt phenomena? an exploration of the law of closure. *ArXiv*, 1903.01069, 2019.

[83] Y. Li, L. Kaiser, S. Bengio, and S. Si. Area attention. In *International Conference on Machine Learning, ICML*, 2019.

[84] M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio. Transfusion: Understanding transfer learning with applications to medical imaging. In *Advances In Neural Information Processing Systems, NeurIPS*, 2019.

[85] J. Chorowski, R. J. Weiss, R. A. Saurous, and S. Bengio. On using backpropagation for speech texture generation and voice conversion. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, 2018.

[86] G. F. Elsayed, D. Krishnan, H. Mobahi, K. Regan, and S. Bengio. Large margin deep networks for classification. In *Advances In Neural Information Processing Systems, NeurIPS*, 2018.

[87] L. Kaiser and S. Bengio. Discrete autoencoders for sequence models. *ArXiv*, 1801.09797, 2018.

[88] L. Kaiser, A. Roy, A. Vaswani, N. Parmar, S. Bengio, J. Uszkoreit, and N. Shazeer. Fast decoding in sequence models using discrete latent variables. In *International Conference on Machine Learning, ICML*, 2018.

[89] A. Kurakin, I. Goodfellow, S. Bengio, Y. Dong, F. Liao, M. Liang, T. Pang, J. Zhu, X. Hu, C. Xie, J. Wang, Z. Zhang, Z. Ren, A. Yuille, S. Huang, Y. Zhao, Y. Zhao, Z. Han, J. Long, Y. Berdibekov, T. Akiba, S. Tokui, and M. Abe. Adversarial attacks and defences competition. *ArXiv*, 1804.00097, 2018.

[90] Y. Li, S. Bengio, and G. Bailly. Predicting human performance in vertical menu selection using deep learning. In *ACM CHI Conference*, 2018.

[91] Y. Li, N. Du, and S. Bengio. Time-dependent representation for neural event sequence prediction. In *Workshop Track of the International Conference on Learning Representations, ICLR*, 2018.

[92] L. Logeswaran, H. Lee, and S. Bengio. Content preserving text generation with attribute controls. In *Advances In Neural Information Processing Systems, NeurIPS*, 2018.

[93] A. S. Morcos, M. Raghu, and S. Bengio. Insights on representational similarity in neural networks with canonical correlation. In *Advances In Neural Information Processing Systems, NeurIPS*, 2018.

[94] A. Vaswani, S. Bengio, E. Brevdo, F. Chollet, A. N. Gomez, S. Gouws, L. Jones, L. Kaiser, N. Kalchbrenner, N. Parmar, R. Sepassi, N. Shazeer, and J. Uskoreit. Tensor2tensor for neural machine transalation. *ArXiv*, 1803.07416, 2018.

[95] C. Zhang, O. Vinyals, R. Munos, and S. Bengio. A study on overfitting in deep reinforcement learning. *ArXiv*, 1804.06893, 2018.

[96] I. Bello, H. Pham, Q. V. Le, M. Norouzi, and S. Bengio. Neural combinatorial optimization with reinforcement learning. In *Workshop Track of the International Conference on Learning Representations, ICLR*, 2017.

[97] C. Chelba, M. Norouzi, and S. Bengio. N-gram language modeling using recurrent neural network estimation. *ArXiv*, 1703.10724, 2017.

[98] L. Dinh, R. Pascanu, S. Bengio, and Y. Bengio. Sharp minima can generalize for deep nets. In *International Conference on Machine Learning, ICML*, 2017.

[99] L. Dinh, J. Sohl-Dickstein, and S. Bengio. Density estimation using real NVP. In *International Conference on Learning Representations, ICLR*, 2017.

[100] L. Kaiser, O. Nachum, A. Roy, and S. Bengio. Learning to remember rare events. In *International Conference on Learning Representations, ICLR*, 2017.

[101] A. Kurakin, I. Goodfellow, and S. Bengio. Adversarial examples in the physical world. In *Workshop Track of the International Conference on Learning Representations, ICLR*, 2017.

[102] A. Kurakin, I. Goodfellow, and S. Bengio. Adversarial machine learning at scale. In *International Conference on Learning Representations, ICLR*, 2017.

[103] A. Mirhoseini, H. Pham, Q. V. Le, B. Steiner, R. Larsen, Y. Zhou, N. Kumar, M. Norouzi, S. Bengio, and J. Dean. Device placement optimization with reinforcement learning. In *International Conference on Machine Learning, ICML*, 2017.

[104] R. Vedantam, S. Bengio, K. Murphy, D. Parikh, and G. Chechik. Context-aware captions from context-agnostic supervision. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2017.

[105] Y. Wang, R.J. Skerry-Ryan, D. Stanton, Y. Wu, R.J. Weiss, N. Jaitly, Z. Yang, Y. Xiao, Z. Chen, S. Bengio, Q. Le, Y. Agiomyrgiannakis, R. Clark, and R.A. Saurous. Tacotron: A fully end-to-end text-to-speech synthesis model. In *Proceedings of Interspeech*, 2017.

[106] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals. Understanding deep learning requires rethinking generalization. In *International Conference on Learning Representations, ICLR*, 2017. Best Paper Award.

[107] S. R. Bowman, L. Vilnis, O. Vinyals, A. M. Dai, R. Jozefowicz, and S. Bengio. Generating sentences from a continuous space. In *SIGNLL Conference on Computational Natural Language Learning, CONLL*, 2016.

[108] J. Chen, R. Monga, S. Bengio, and R. Jozefowicz. Revisiting distributed synchronous SGD. In *Workshop Track of the International Conference on Learning Representations, ICLR*, 2016.

[109] M. Cissé, M. Al-Shedivat, and S. Bengio. ADIOS: Architectures deep in output space. In *Proceedings of the 33rd International Conference on Machine Learning, ICML*, 2016.

[110] G. Heigold, I. Moreno, S. Bengio, and N. Shazeer. End-to-end text-dependent speaker verification. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, 2016.

[111] N. Jaitly, D. Sussillo, Q. V. Le, O. Vinyals, I. Sutskever, and S. Bengio. An online sequence-to-sequence model using partial conditioning. In *Advances In Neural Information Processing Systems, NIPS*, 2016.

[112] L. Kaiser and S. Bengio. Can active memory replace attention? In *Advances In Neural Information Processing Systems, NIPS*, 2016.

[113] M. Norouzi, S. Bengio, Z. Chen, N. Jaitly, M. Schuster, Y. Wu, and D. Schuurmans. Reward augmented maximum likelihood for neural structured prediction. In *Advances In Neural Information Processing Systems, NIPS*, 2016.

[114] O. Vinyals, S. Bengio, and M. Kudlur. Order matters: Sequence to sequence for sets. In *International Conference on Learning Representations, ICLR*, 2016.

[115] S. Bengio, O. Vinyals, N. Jaitly, and N. Shazeer. Scheduled sampling for sequence prediction with recurrent neural networks. In *Advances In Neural Information Processing Systems, NIPS*, 2015.

[116] V. Ramanathan, J. Deng, C. Li, W. Han, Z. Li, K. Gu, Y. Song, S. Bengio, C. Rosenberg, and F.-F. Li. Learning semantic relationships for better action retrieval in images. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2015.

[117] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan. Show and tell: A neural image caption generator. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2015.

[118] S. Bengio and G. Heigold. Word embeddings for speech recognition. In *Proceedings of the 15th Conference of the International Speech Communication Association, Interspeech*, 2014.

[119] J. Deng, N. Ding, Y. Jia, A. Frome, K. Murphy, S. Bengio, Y. Li, H. Neven, and H. Adam. Large-scale object classification using label relation graphs. In *Proceedings of the European Conference on Computer Vision, ECCV*, 2014. Best Paper Award.

[120] I. Lapidot, J.-F. Bonastre, and S. Bengio. Telephone conversation speaker diarization using mealy-HMMs. In *Proceedings of Speaker Odyssey*, 2014.

[121] J. Lee, S. Bengio, S. Kim, G. Lebanon, and Y. Singer. Local collaborative ranking. In *International World Wide Web Conference, WWW*, 2014.

[122] M. Norouzi, T. Mikolov, S. Bengio, Y. Singer, J. Shlens, A. Frome, G. S. Corrado, and J. Dean. Zero-shot learning by convex combination of semantic embeddings. In *International Conference on Learning Representations, ICLR*, 2014.

[123] S. Bengio, J. Dean, D. Erhan, E. Ie, Q. Le, A. Rabinovich, J. Shlens, and Y. Singer. Using web co-occurrence statistics for improving image categorization. *ArXiv*, 1312.5697, 2013.

[124] A. Frome, G. Corrado, J. Shlens, S. Bengio, J. Dean, M. Ranzato, and T. Mikolov. DeViSE: A deep visual-semantic embedding model. In *Advances In Neural Information Processing Systems, NIPS*, 2013.

[125] S. Bengio. Large scale visual semantic extraction. In *Frontiers of Engineering - Reports on Leading-Edge Engineering from the 2011 Symposium*, 2012.

[126] J. Weston, S. Bengio, and N. Usunier. Wsabie: Scaling up to large vocabulary image annotation. In *Proceedings of the International Joint Conference on Artificial Intelligence, IJCAI*, 2011.

[127] S. Bengio, J. Weston, and D. Grangier. Label embedding trees for large multi-class tasks. In *Advances in Neural Information Processing Systems, NIPS*, 2010.

[128] J. Weston, S. Bengio, and N. Usunier. Large scale image annotation: Learning to rank with joint word-image embeddings. In *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, ECML-PKDD*, 2010. Best Paper Award in Machine Learning.

[129] S. Bengio, F. Pereira, Y. Singer, and D. Strelow. Group sparse coding. In *Advances in Neural Information Processing Systems, NIPS*. MIT Press, 2009.

[130] G. Chechik, V. Sharma, U. Shalit, and S. Bengio. Large-scale online learning of image similarity through ranking: Extended abstract. In *4th Iberian Conference on Pattern Recognition and Image Analysis IbPRIA*, 2009.

[131] G. Chechik, V. Sharma, U. Shalit, and S. Bengio. An online algorithm for large scale image similarity learning. In *Advances in Neural Information Processing Systems, NIPS*. MIT Press, 2009.

[132] D. Erhan, P.-A. Manzagol, Y. Bengio, S. Bengio, and P. Vincent. The difficulty of training deep architectures and the effect of unsupervised pre-training. In D. van Dyk and M. Wellings, editors, *Proceedings of The Twelfth International Conference on Artificial Intelligence and Statistics, AISTATS*, volume 5 of *JMLR Workshop and Conference Procedings*, pages 153–160, 2009.

[133] M. Rehn, R. F. Lyon, S. Bengio, T. C. Walters, and G. Chechik. Sound ranking using auditory sparse-code representations. In *ICML 2009 Workshop on Sparse Method for Music Audio*, 2009.

[134] G. Chechik, E. Ie, M. Rehn, S. Bengio, and D. Lyon. Large-scale content-based audio retrieval from text queries. In *ACM International Conference on Multimedia Information Retrieval, MIR*, 2008.

[135] J.-F. Paiement, Y. Grandvalet, S. Bengio, and D. Eck. A distance model for rhythms. In *International Conference on Machine Learning, ICML*, 2008.

[136] S. Bengio and J. Mariéthoz. Biometric person authentication is a multiple classifier problem. In M. Haindl, J. Kittler, and F. Roli, editors, *7th International Workshop on Multiple Classifier Systems, MCS, Lecture Notes in Computer Science*, volume LNCS 4472. Springer-Verlag, 2007.

[137] D. Grangier and S. Bengio. Learning the inter-frame distance for discriminative template-based keyword detection. In *Proceedings of the 10th European Conference on Speech Communication and Technology, Eurospeech-Interspeech*, 2007.

[138] J. Keshet, D. Grangier, and S. Bengio. Discriminative keyword spotting. In *ISCA Research Workshop on Non Linear Speech Processing, NOLISP*, 2007.

[139] J.-F. Paiement, Y. Grandvalet, S. Bengio, and D. Eck. A generative model for rhythms. In *NIPS Workshop on Brain, Music and Cognition*, 2007.

[140] H. Paugam-Moisy, R. Martinez, and S. Bengio. A supervised learning approach based on STDP and polychronization in spiking neuron networks. In *European Symposium on Artificial Neural Networks, ESANN*, 2007.

[141] N. Poh and S. Bengio. Estimating the confidence interval of expected performance curve in biometric authentication using joint bootstrap. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, 2007.

[142] D. Zhang and S. Bengio. Exploring contextual information in a layered framework for group action recognition. In *IEEE International Conference on Multimedia & Expo, ICME*, 2007.

[143] O. Glickman, I. Dagan, M. Keller, S. Bengio, and W. Daelemans. Investigating lexical substitution scoring for subtitle generation. In *Tenth Conference on Computational Natural Language Learning, CONLL*, 2006.

[144] D. Grangier and S. Bengio. A neural network to retrieve images from text queries. In *Proceedings of the 16th International Conference on Artificial Neural Networks: Biological Inspirations, ICANN, Lecture Notes in Computer Science*, volume LNCS 4132. Springer-Verlag, 2006.

[145] D. Grangier, F. Monay, and S. Bengio. A discriminative approach for the retrieval of images from text queries. In *European Conference on Machine Learning, ECML, Lecture Notes in Computer Science*, volume LNCS 4212. Springer-Verlag, 2006.

[146] D. Grangier, F. Monay, and S. Bengio. Learning to retrieve images from text queries with a discriminative model. In *International Workshop on Adaptive Multimedia Retrieval, AMR*, 2006.

[147] J. Keshet, S. Shalev-Shwartz, S. Bengio, Y. Singer, and D. Chazan. Discriminative kernel-based phoneme sequence recognition. In *Proceedings of the International Conference on Spoken Language Processing, Interspeech-ICSLP*, 2006.

[148] H. Ketabdar, J. Vepa, S. Bengio, and H. Bourlard. Posterior based keyword spotting with a priori thresholds. In *Proceedings of the International Conference on Spoken Language Processing, Interspeech-ICSLP*, 2006.

[149] H. Ketabdar, J. Vepa, S. Bengio, and H. Bourlard. Using more informative posterior probabilities for speech recognition. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, 2006.

[150] M. Liwicki, A. Schlapbach, H. Bunke, S. Bengio, J. MariÃ©thoz, and J. Richiardi. Writer identification for smart meeting room systems. In H. Bunke and A. L. Spitz, editors, *Document Analysis Systems VII: 7th International Workshop, DAS, Lecture Notes in Computer Science*, volume LNCS 3872, pages 186–195. Springer-Verlag, 2006.

[151] J. MariÃ©thoz and S. Bengio. A max kernel for text-independent speaker verification systems. In *Second Workshop on Multimodal User Authentication, MMUA*, 2006.

[152] J.-F. Paiement, D. Eck, and S. Bengio. Probabilistic melodic harmonization. In L. Lamontagne and M. Marchand, editors, *Advances in Artificial Intelligence: 19th Conference of the Canadian Society for Computational Studies of Intelligence, Canadian AI, Lecture Notes in Computer Science*, volume LNCS 4013, pages 218–229. Springer-Verlag, 2006.

[153] N. Poh and S. Bengio. Chimeric users to construct fusion classifiers in biometric authentication tasks: An investigation. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, 2006.

[154] N. Poh, S. Bengio, and A. Ross. Revisiting Doddington's zoo: A systematic method to assess user-dependent variabilities. In *Second Workshop on Multimodal User Authentication, MMUA*, 2006.

[155] A. Pozdnoukhov and S. Bengio. Graph-based invariant manifolds for invariant pattern recognition with kernel methods. In *International Conference on Pattern Recognition, ICPR*, 2006.

[156] A. Pozdnoukhov and S. Bengio. Semi-supervised kernel methods for regression estimation. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, 2006.

[157] D. Zhang, D. Gatica-Perez, D. Roy, and S. Bengio. Modeling interactions from email communication. In *IEEE International Conference on Multimedia & Expo, ICME*, 2006.

[158] S. Bengio and H. Bourlard. Multi channel sequence processing. In J. Winkler, M. Niranjan, and N. Lawrence, editors, *Deterministic and Statistical Methods in Machine Learning: First International Workshop, Lecture Notes in Artificial Intelligence*, volume LNAI 3635, pages 22–36. Springer-Verlag, 2005.

[159] S. Bengio, J. MariÃ©thoz, and M. Keller. The expected performance curve. In *International Conference on Machine Learning, ICML, Workshop on ROC Analysis in Machine Learning*, 2005.

[160] C. Dimitrakakis and S. Bengio. Boosting word error rates. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, pages 501–504, 2005.

[161] C. Dimitrakakis and S. Bengio. Gradient-based estimates of return distributions. In *PASCAL Workshop on Principled Methods of Trading Exploration and Exploitation*, 2005.

[162] D. Gatica-Perez, D. Zhang, and S. Bengio. Extracting information from multimedia meeting collections. In *7th ACM SIGMM International Workshop on Multimedia Information Retrieval, MIR*, 2005.

[163] D. Gatica-Perez, I. McCowan D. Zhang, and S. Bengio. Detecting group interest-level in meetings. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, pages 489–492, 2005.

[164] Y. Grandvalet, J. MariÃ©thoz, and S. Bengio. A probabilistic interpretation of SVMs with an application to unbalanced classification. In *Advances in Neural Information Processing Systems, NIPS 18*. MIT Press, 2005.

[165] D. Grangier and S. Bengio. Exploiting hyperlinks to learn a retrieval model. In *Proceedings of the NIPS 2005 Workshop on Learning to Rank*, 2005.

[166] D. Grangier and S. Bengio. Inferring document similarity from hyperlinks. In *Proceedings of the Conference on Information and Knowledge Management, CIKM*, 2005.

[167] M. Keller and S. Bengio. A neural network for text representation. In *Proceedings of the 15th International Conference on Artificial Neural Networks: Biological Inspirations, ICANN, Lecture Notes in Computer Science*, volume LNCS 3697, pages 667–672. Springer-Verlag, 2005.

[168] M. Keller, S. Bengio, and S. Y. Wong. Benchmarking non-parametric statistical tests. In *Advances in Neural Information Processing Systems, NIPS 18*. MIT Press, 2005.

[169] H. Ketabdar, H. Bourlard, and S. Bengio. Hierarchical multi-stream posterior based speech recognition system. In *Machine Learning for Multimodal Interactions: Second International Workshop, MLMI, Lecture Notes in Computer Science*, volume LNCS 3869, 2005.

[170] H. Ketabdar, J. Vepa, S. Bengio, and H. Bourlard. Developing and enhancing posterior based speech recognition systems. In *Proceedings of the 9th European Conference on Speech Communication and Technology, Eurospeech-Interspeech*, 2005.

[171] J.-F. Paiement, D. Eck, and S. Bengio. A probabilistic model for chord progressions. In *International Conference on Music Information Retrieval, ISMIR*, 2005.

[172] J.-F. Paiement, D. Eck, S. Bengio, and D. Barber. A graphical model for chord progressions embedded in a psychoacoustic space. In *International Conference on Machine Learning, ICML*, 2005.

[173] N. Poh and S. Bengio. Can chimeric persons be used in multimodal biometric authentication experiments? In S. Renals and S. Bengio, editors, *Machine Learning for Multimodal Interactions: Second International Workshop, MLMI*, volume LNCS 3869. Springer-Verlag, 2005.

[174] N. Poh and S. Bengio. EER of fixed and trainable fusion classifiers: A theoretical study with application to biometric authentication tasks. In N. C. Oza, R. Polikar, and J. Kittler, editors, *6th International Workshop on Multiple Classifier Systems, MCS, Lecture Notes in Computer Science*, volume LNCS 3541, pages 74–85. Springer-Verlag, 2005.

[175] N. Poh and S. Bengio. F-ratio client-dependent normalisation for biometric authentication tasks. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, pages 721–724, 2005.

[176] N. Poh and S. Bengio. Improving fusion with margin-derived confidence in biometric authentication tasks. In T. Kanade, A. Jain, and N. K. Ratha, editors, *5th International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA, Lecture Notes in Computer Science*, volume LNCS 3546, pages 1059–1068. Springer-Verlag, 2005.

[177] N. Poh and S. Bengio. A novel approach to combining client-dependent and confidence information in multimodal biometrics. In T. Kanade, A. Jain, and N. K. Ratha, editors, *5th International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA, Lecture Notes in Computer Science*, volume LNCS 3546, pages 1120–1129. Springer-Verlag, 2005.

[178] N. Poh and S. Bengio. A score-level fusion benchmark database for biometric authentication. In T. Kanade, A. Jain, and N. K. Ratha, editors, *5th International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA, Lecture Notes in Computer Science*, volume LNCS 3546, pages 474–483. Springer-Verlag, 2005.

[179] A. Pozdnoukhov and S. Bengio. Improving kernel classifiers for object categorization problems. In *International Conference on Machine Learning, ICML, Workshop on Learning with Partially Classified Training Data*, 2005.

[180] C. Sanderson, F. Cardinaux, and S. Bengio. On accuracy/robustness/complexity trade-offs in face verification. In *IEEE International Conference on Information Technology and Applications, ICITA*, pages 638–643, 2005.

[181] D. Zhang, D. Gatica-Perez, S. Bengio, and I. McCowan. Semi-supervised adapted HMMs for unusual event detection. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2005.

[182] D. Zhang, D. Gatica-Perez, S. Bengio, and I. McCowan. Semi-supervised meeting event recognition with adapted HMMs. In *IEEE International Conference on Multimedia Expo, ICME*, pages 611–618, 2005.

[183] D. Zhang, D. Gatica-Perez, S. Bengio, and D. Roy. Learning influence among interacting markov chains. In *Advances in Neural Information Processing Systems, NIPS 18*. MIT Press, 2005.

[184] S. Bengio and J. MariÃ©thoz. The expected performance curve: a new assessment measure for person authentication. In *Proceedings of Odyssey 2004: The Speaker and Language Recognition Workshop*, 2004.

[185] S. Bengio and J. MariÃ©thoz. A statistical significance test for person authentication. In *Proceedings of Odyssey 2004: The Speaker and Language Recognition Workshop*, 2004.

[186] H. Bourlard, S. Bengio, M. Magimai Doss, Q. Zhu, B. Mesot, and N. Morgan. Towards using hierarchical posteriors for flexible automatic speech recognition systems. In *Proceedings of the DARPA EARS (Effective, Affordable, Reusable, Speech-to-text) Rich Transcription (RT'04) Workshop*, 2004.

[187] F. Cardinaux, C. Sanderson, and S. Bengio. Face verification using adapted generative models. In *International Conference on Automatic Face and Gesture Recognition, FG*, pages 825–830, 2004.

[188] S. Chiappa and S. Bengio. HMM and IOHMM modeling of EEG rhythms for asynchronous BCI systems. In *European Symposium on Artificial Neural Networks, ESANN*, 2004.

[189] R. Collobert and S. Bengio. A gentle hessian for efficient gradient descent. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, volume 5, pages 517–520, 2004.

[190] R. Collobert and S. Bengio. Links between perceptrons, MLPs and SVMs. In *International Conference on Machine Learning, ICML*, 2004.

[191] C. Dimitrakakis and S. Bengio. Boosting HMMs with an application to speech recognition. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, volume 5, pages 621–624, 2004.

[192] C. Dimitrakakis and S. Bengio. Online policy adaptation for ensemble classifiers. In *European Symposium on Artificial Neural Networks, ESANN*, 2004.

[193] M. Magimai Doss, S. Bengio, and H. Bourlard. Joint decoding for phoneme-grapheme continuous speech recognition. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, volume 1, pages 177–180, 2004.

[194] M. Keller and S. Bengio. Theme topic mixture model: A graphical model for document representation. In *PASCAL Workshop on Learning Methods for Text Understanding and Mining*, 2004.

[195] I. McCowan, D. Gatica-Perez, S. Bengio, D. Moore, and H. Bourlard. Towards computer understanding of human interactions. In *Machine Learning for Multimodal Interaction: First International Workshop, MLMI, Lecture Notes in Computer Science*, volume LNCS 3361, pages 56–75. Springer-Verlag, 2004.

[196] K. Messer, J. Kittler, M. Sadeghi, M. Hamouz, A. Kostin, F. Cardinaux, S. Marcel, S. Bengio, C. Sanderson, N. Poh, Y. Rodriguez, J. Czyz, L. Vandendorpe, C. McCool, S. Lowther, S. Sridharan, V. Chandran, R. Paredes, E. Vidal, L. Bai, L. Shen, Y. Wang, C. Yueh-Hsuan, L. Hsien-Chang, H. Yi-Ping, A. Heinrichs, M. Muller, A. Tewes, C. von der Malsburg, R. Wurtz, Z. Wang, F. Xue, Y. Ma, Q. Yang, C. Fang, X. Ding, S. Lucey, R. Goss, and H. Schneiderman. Face authentication test on the BANCA database. In *International Conference on Pattern Recognition, ICPR*, volume 4, pages 523–532, 2004.

[197] K. Messer, J. Kittler, M. Sadeghi, M. Hamouz, A. Kostin, S. Marcel, S. Bengio, F. Cardinaux, C. Sanderson, N. Poh, Y. Rodriguez, K. Kryszczuk, J. Czyz, L. Vandendorpe, J. Ng, H. Cheung, and B. Tang. Face authentication competition on the BANCA database. In *International Conference on Biometric Authentication, ICBA, Lecture Notes in Computer Science*, volume LNCS 3072, pages 8–15. Springer-Verlag, 2004.

[198] N. Poh and S. Bengio. Noise-robust multi-stream fusion for text-independent speaker authentication. In *Proceedings of Odyssey 2004: The Speaker and Language Recognition Workshop*, 2004.

[199] N. Poh and S. Bengio. Towards predicting optimal subsets of base classifiers in biometric authentication tasks. In S. Bengio and H. Bourlard, editors, *Machine Learning for Multimodal Interactions: First International Workshop, MLMI, Lecture Notes in Computer Science*, volume LNCS 3361, pages 159–172. Springer-Verlag, 2004.

[200] N. Poh and S. Bengio. Why do multi-stream, multi-band and multi-modal approaches work on biometric user authentication tasks? In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, volume 5, pages 893–896, 2004.

[201] N. Poh, C. Sanderson, and S. Bengio. Spectral subband centroids as complementary features for speaker authentication. In *International Conference on Biometric Authentication, ICBA, Lecture Notes in Computer Science*, volume LNCS 3072, pages 631–639. Springer-Verlag, 2004.

[202] A. Pozdnoukhov and S. Bengio. Tangent vector kernels for invariant image classification with SVMs. In *International Conference on Pattern Recognition, ICPR*, volume 3, pages 486–489, 2004.

[203] Y. Rodriguez, F. Cardinaux, S. Bengio, and J. MariÃ©thoz. Estimating the quality of face localization for face verification. In *IEEE International Conference on Image Processing, ICIP*, pages 581–584, 2004.

[204] C. Sanderson and S. Bengio. Extrapolating single view face models for multi-view recognition. In *International Conference on Intelligente Sensors, Sensor Networks and Information Processings, ISSNIP*, pages 581–586, 2004.

[205] C. Sanderson and S. Bengio. Statistical transformations of frontal models for non-frontal face verification. In *IEEE International Conference on Image Processing, ICIP*, pages 585–588, 2004.

[206] D. Zhang, D. Gatica-Perez, S. Bengio, I. McCowan, and G. Lathoud. Modeling individual and group actions in meetings: a two-layer hmm framework. In *IEEE Workshop on Event Mining at the Conference on Computer Vision and Pattern Recognition, CVPR*, 2004.

[207] D. Zhang, D. Gatica-Perez, S. Bengio, I. McCowan, and G. Lathoud. Multimodal group action clustering in meetings. In *ACM Multimedia Workshop on Video Surveillance and Sensor Networks*, 2004.

[208] E. Bailly-BailliÃ¨re, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. MariÃ©thoz, J. Matas, K. Messer, V. Popovici, F. PorÃ©e, B. Ruiz, and J.-P. Thiran. The BANCA database and evaluation protocol. In *4th International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA, Lecture Notes in Computer Science*, volume LNCS 2688, pages 625–638. Springer-Verlag, 2003.

[209] M. Barnard, J.-M. Odobez, and S. Bengio. Multi-modal audio-visual event recognition for football analysis. In *IEEE Workshop on Neural Networks for Signal Processing, NNSP*, pages 469–478, 2003.

[210] S. Bengio. An asynchronous hidden markov model for audio-visual speech recognition. In S. Becker, S. Thrun, and K. Obermayer, editors, *Advances in Neural Information Processing Systems, NIPS 15*, pages 1237–1244. MIT Press, 2003.

[211] S. Bengio. Multimodal authentication using asynchronous HMMs. In *4th International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA, Lecture Notes in Computer Science*, volume LNCS 2688, pages 770–777. Springer-Verlag, 2003.

[212] J. Czyz, S. Bengio, C. Marcel, and L. Vandendorpe. Scalability analysis of audio-visual person identity verification. In *4th International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA, Lecture Notes in Computer Science*, volume LNCS 2688, pages 752–760. Springer-Verlag, 2003.

[213] M. Magimai Doss, T. A. Stephenson, H. Bourlard, and S. Bengio. Phoneme-grapheme based speech recognition system. In *IEEE Automatic Speech Recognition and Understanding Workshop, ASRU*, pages 94–98, 2003.

[214] D. Gatica-Perez, I. McCowan, M. Barnard, S. Bengio, and H. Bourlard. On automatic annotation of meeting databases. In *IEEE International Conference on Image Processing, ICIP*, volume 3, pages 629–632, 2003.

[215] Q. Le and S. Bengio. Client dependent GMM-SVM models for speaker verification. In *International Conference on Artificial Neural Networks, ICANN/ICONIP, Lecture Notes in Computer Science*, volume LNCS 2714, pages 443–451. Springer Verlag, 2003.

[216] I. McCowan, S. Bengio, D. Gatica-Perez, G. Lathoud, F. Monay, D. Moore, P. Wellner, and H. Bourlard. Modeling human interaction in meetings. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, volume 4, pages 748–751, 2003.

[217] I. McCowan, D. Gatica-Perez, S. Bengio, D. Moore, and H. Bourlard. Towards computer understanding of human interactions. In *Ambient Intelligence, Lecture Notes in Computer Science*, volume LNCS 2875, pages 235–251, Eindhoven, 2003. Springer-Verlag.

[218] K. Messer, J. Kittler, M. Sadeghi, S. Marcel, C. Marcel, S. Bengio, F. Cardinaux, C. Sanderson, J. Czyz, L. Vandendorpe, S. Srisuk, M. Petrou, W. Kurutach, A. Kadyrov, R. Paredes, B. Kepenekci, F. B. Tek, G. B. Akar, F. Deravi, and N. Mavity. Face verification competition on the XM2VTS database. In *4th International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA, Lecture Notes in Computer Science*, volume LNCS 2688, pages 964–974. Springer-Verlag, 2003.

[219] N. Poh and S. Bengio. Non-linear variance reduction techniques in biometric authentication. In *IEEE Multimodal User Authentication Workshop*, 2003.

[220] N. Poh, S. Marcel, and S. Bengio. Improving face authentication using virtual samples. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, volume 3, pages 233–236, 2003.

[221] C. Sanderson and S. Bengio. Augmenting frontal face models for non-frontal verification. In *IEEE Multimodal User Authentication Workshop*, 2003.

[222] C. Sanderson and S. Bengio. Robust features for frontal authentication in difficult image conditions. In *4th International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA, Lecture Notes in Computer Science*, volume LNCS 2688, pages 495–504. Springer-Verlag, 2003.

[223] C. Sanderson, S. Bengio, H. Bourlard, J. MariÃ©thoz, R. Collobert, M.F. BenZeghiba, F. Cardinaux, and S. Marcel. Speech & face based biometric authentication at IDIAP. In *International Conference on Multimedia and Expo, ICME*, volume 3, pages 1–4, 2003.

[224] A. Vinciarelli, S. Bengio, and H. Bunke. Offline recognition of large vocabulary cursive handwritten text. In *International Conference on Document Analysis and Recognition, ICDAR*, pages 1101–1105, 2003.

[225] R. Collobert, S. Bengio, and Y. Bengio. A parallel mixture of SVMs for very large scale problems. In T.G. Dietterich, S. Becker, and Z. Ghahramani, editors, *Advances in Neural Information Processing Systems, NIPS 14*, pages 633–640. MIT Press, 2002.

[226] R. Collobert, S. Bengio, and J. MariÃ©thoz. Torch: a modular machine learning software library. *IDIAP RR*, 02-46, 2002.

[227] R. Collobert, Y. Bengio, and S. Bengio. Scaling large learning problems with hard parallel mixtures. In S. Lee and A. Verri, editors, *International Workshop on Pattern Recognition with Support Vector Machines, SVM, Lecture Notes in Computer Science*, volume LNCS 2388, pages 8–23. Springer-Verlag, 2002.

[228] N. Gilardi, S. Bengio, and M. Kanevski. Conditional gaussian mixture models for environmental risk mapping. In *IEEE Workshop on Neural Networks for Signal Processing, NNSP*, pages 777–786, 2002.

[229] S. Marcel and S. Bengio. Improving face verification using skin color information. In *Proceedings of the 16th International Conference on Pattern Recognition, ICPR*, volume 2, pages 11–15. IEEE Computer Society Press, 2002.

[230] S. Marcel, C. Marcel, and S. Bengio. A state-of-the-art neural network for robust face verification. In *COST275 Workshop on the advent of Biometrics on the Internet*, 2002.

[231] J. MariÃ©thoz and S. Bengio. A comparative study of adaptation methods for speaker verification. In *Proceedings of the International Conference on Spoken Language Processing, ICSLP*, 2002.

[232] N. Poh, S. Bengio, and J. Korczak. A multi-sample multi-source model for biometric authentication. In *IEEE Workshop on Neural Networks for Signal Processing, NNSP*, pages 375–384, 2002.

[233] A. Vinciarelli and S. Bengio. Offline cursive word recognition using continuous density hidden markov models trained with PCA or ICA features. In *Proceedings of the 16th International Conference on Pattern Recognition, ICPR*, volume 3, pages 81–84. IEEE Computer Society Press, 2002.

[234] A. Vinciarelli and S. Bengio. Writer adaptation techniques in HMM based off-line cursive script recognition. In *Proceedings of the 8th International Conference on Frontiers in Handwriting Recognition*, pages 287–291, 2002.

[235] K. Weber, S. Bengio, and H. Bourlard. Increasing speech recognition noise robustness with HMM2. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, volume 1, pages 929–932, 2002.

[236] K. Weber, F. de Wet, B. Cranen, L. Boves, S. Bengio, and H. Bourlard. Evaluation of formant-like features for ASR. In *Proceedings of the International Conference on Spoken Language Processing, ICSLP*, 2002.

[237] S. Bengio and J. MariÃ©thoz. Learning the decision function for speaker verification. In *IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, volume 1, pages 425–428, 2001.

[238] H. Bourlard, S. Bengio, and K. Weber. New approaches towards robust and adaptive speech recognition. In T.K. Leen, T.G. Dietterich, and V. Tresp, editors, *Advances in Neural Information Processing Systems, NIPS 13*, pages 751–757. MIT Press, 2001.

[239] J.-L. DesGranges, P. Agin, and S. Bengio. The use of predictive models of breeding bird assemblages for assessing and monitoring forest bird diversity. In A. Franc, O. Laroussinie, and T. Karjalainen, editors, *Criteria and Indicators for Sustainable Forest Management at the Forest Management Unit Level*, volume 38, pages 181–200. European Forest Institute Proceedings, 2001.

[240] K. Weber, S. Bengio, and H. Bourlard. HMM2- extraction of formant features and their use for robust ASR. In *Proceedings of the European Conference on Speech Communication and Technology, EUROSPEECH*, 2001.

[241] K. Weber, S. Bengio, and H. Bourlard. Speech recognition using advanced HMM2 features. In *Proceedings of the Automatic Speech Recognition and Understanding Workshop, ASRU*, pages 65–68, 2001.

[242] Y. Bengio and S. Bengio. Modeling high-dimensional discrete data with multi-layer neural networks. In S.A. Solla, T.K. Leen, and K.-R. MÃ¼ller, editors, *Advances in Neural Information Processing Systems, NIPS 12*, pages 400–406. MIT Press, 2000.

[243] T. A. Stephenson, H. Bourlard, S. Bengio, and A. C. Morris. Automatic speech recognition using dynamic Bayesian networks with both acoustic and articulatory variables. In *Proceedings of the International Conference on Speech and Language Processing, ICSLP*, Beijing, China, October 2000.

[244] K. Weber, S. Bengio, and H. Bourlard. HMM2- a novel approach to HMM emission probability estimation. In *Proceedings of the International Conference on Speech and Language Processing, ICSLP*, Beijing, China, October 2000.

[245] Y. Bengio, S. Bengio, J.-F. Isabelle, and Y. Singer. Shared context probabilistic transducers. In *Advances in Neural Information Processing Systems, NIPS 10*, 1998.

[246] S. Bengio, F. Clerot, A. Gravey, and D. Collobert. Dynamical resource reservation schemes in an ATM network using neural network-based traffic prediction. In D. D. Kouvatsos, editor, *Proceedings of the IFIP Fifth International Workshop on Performance Modelling and Evaluation of ATM Networks*. Kluwer B. V., 1997.

[247] S. Bengio and Y. Bengio. An EM algorithm for asynchronous input/output hidden markov models. In *Proceedings of the International Conference on Neural Information Processing, ICONIP*, Hong Kong, 1996.

[248] Y. Bengio and S. Bengio. Training asynchronous input/output hidden markov models. In *AAAI Spring Symposium on Computational Issues in Learning Models of Dynamical Systems*, 1996.

[249] S. Bengio, F. Fessant, and D. Collobert. A connectionist system for medium-term horizon time series prediction. In *International Workshop on Applications of Neural Networks to Telecommunications, IWANNT*, Stockholm, Sweden, 1995.

[250] S. Bengio, Y. Bengio, and J. Cloutier. Use of genetic programming for the search of a new learning rule for neural networks. In *Proceedings of the First Conference on Evolutionary Computation, IEEE World Congress on Computational Intelligence*, volume 1, pages 324–327, 1994.

[251] S. Bengio, Y. Bengio, J. Cloutier, and J. Gecsei. Generalization of a parametric learning rule. In S. Gielen and B. Kappen, editors, *Proceedings of the International Conference on Artificial Neural Networks, ICANN'93*, pages 502–502, Amsterdam, Nederlands, 1993. Springer-Verlag.

[252] S. Bengio, Y. Bengio, J. Cloutier, and J. Gecsei. Aspects thÃ©oriques de l'optimisation d'une rÃ¨gle d'apprentissage. In *Actes de la confÃ©rence Neuro-NÃ®mes 1992*, NÃ®mes, France, 1992.

[253] S. Bengio, Y. Bengio, J. Cloutier, and J. Gecsei. On the optimization of a synaptic learning rule. In *Conference on Optimality in Biological and Artificial Networks*, Dallas, USA, 1992.

[254] Y. Bengio, S. Bengio, and J. Cloutier. Learning a synaptic learning rule. In *Proceedings of the International Joint Conference on Neural Networks, IJCNN*, volume 2, pages 969–974, Seattle, USA, 1991.

[255] Y. Bengio, S. Bengio, Y. Pouliot, and P. Agin. A neural network to detect homologies in proteins. In *Advances in Neural Information Processing Systems, NIPS 2*, San Mateo, CA, USA, 1990. Morgan Kaufmann.

[256] S. Bengio, C. Frasson, and J. Gecsei. Integrating traditional and intelligent computerized tutoring. In *Fourth International Symposium on Computer and Information Sciences*, Cesme, Turkey, 1989.

[257] S. Bengio, C. Frasson, and J. Gecsei. Utilisation de systÃ¨mes d'EAO dans des systÃ¨mes d'EIAO. In *SixiÃ¨me symposium canadien sur la technologie pÃ©dagogique*, Halifax, Canada, 1989.

[258] Y. Desmedt, C. Goutier, and S. Bengio. Special uses and abuses of the Fiat-Shamir passport protocol. In *Advances in Cryptology, Crypto, Lecture Notes in Computer Science*, volume LNCS 293, pages 21–39, Santa Barbara, USA, 1988. Springer Verlag.

Edited Books, Book Chapters and Theses

[259] S. Escalera, M. Weimer, M. Burtsev, V. Malykh, V. Logacheva, R. Lowe, I. V. Serban, Y. Bengio, A. Rudnicky, A. W. Black, S. Prabhumoye, Â¿. KidziÂ¿ski, S. P. Mohanty, C. F. Ong, J. L. Hicks, S. Levine, M. SalathÃ©, S. Delp, I. Huerga, A. Grigorenko, L. Thorbergsson, A. D. Nemitz, J. Sandker, S. King, A. S. Ecker, L. A. Gatys, M. Bethge, J. Boyd-Graber, S. Feng, P. Rodriguez, M. Iyyer, H. He, H. DaumÃ© III, S. McGregor, A. Banifatemi, A. Kurakin, I. Goodfellow, and S. Bengio. Introduction to nips 2017 competition track. In S. Escalera and M. Weimer, editors, *The NIPS '17 Competition: Building Intelligent Systems*. Springer, 2018.

[260] M. Stevens, S. Bengio, and Y. Singer. Efficient learning of sparse ranking functions. In B. Scholkopf, Z. Luo, and V. Vovk, editors, *Empirical Inference*. Springer, 2013.

[261] S. Bengio. Statistical machine learning for HCI. In J.-P. Thiran, F. Marqués, and H. Bourlard, editors, *Multimodal Signal Processing: Theory and Applications for Human-Computer Interaction*, pages 7–23. Academic Press, 2010.

[262] S. Bengio and J. Keshet. Introduction. In J. Keshet and S. Bengio, editors, *Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods*, pages 3–10. Wiley, 2009.

[263] D. Grangier, J. Keshet, and S. Bengio. Discriminative keyword spotting. In J. Keshet and S. Bengio, editors, *Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods*, pages 175–194. Wiley, 2009.

[264] J. Keshet and S. Bengio, editors. *Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods*. Wiley, 2009.

[265] J. Mariéthoz, S. Bengio, and Y. Grandvalet. Kernel-based text-independent speaker verification. In J. Keshet and S. Bengio, editors, *Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods*, pages 195–220. Wiley, 2009.

[266] S. Renals, S. Bengio, and J. G. Fiscus, editors. *Machine Learning for Multimodal Interaction: Third International Workshop, MLMI'2006*, volume 4299 of *Lecture Notes in Computer Science*. Springer-Verlag, 2007.

[267] S. Renals and S. Bengio, editors. *Machine Learning for Multimodal Interaction: Second International Workshop, MLMI'2005*, volume 3869 of *Lecture Notes in Computer Science*. Springer-Verlag, 2006.

[268] S. Bengio and H. Bourlard, editors. *Machine Learning for Multimodal Interaction: First International Workshop, MLMI'2004*, volume 3361 of *Lecture Notes in Computer Science*. Springer-Verlag, 2005.

[269] N. Gilardi and S. Bengio. Machine learning for automatic environmental mapping: when and how? In G. Dubois, editor, *Automatic mapping algorithms for routine and emergency monitoring data. Report on the Spatial Interpolation Comparison (SIC2004) exercise*, pages 123–138. Office for Official Publications of the European Communities, Luxembourg, 2005.

[270] H. Bourlard, S. Bengio, and K. Weber. Towards robust and adaptive speech recognition models. In M. Johnson, S. Khudanpur, M. Ostendorf, and R. Rosenfeld, editors, *Mathematical Foundations of Speech and Language Processing*, Institute for Mathematics and its Applications (IMA) Series, Volume 138, pages 169–189. Springer-Verlag, 2003.

[271] N. Gilardi and S. Bengio. Comparison of four machine learning algorithms for spatial data analysis. In G. Dubois, J. Malczewski, and M. DeCort, editors, *Mapping radioactivity in the environment - Spatial Interpolation Comparison 97*, pages 222–237. Office for Official Publications of the European Communities, Luxembourg, 2003.

[272] H. Bourlard, T. Adali, S. Bengio, J. Larsen, and S. Douglas, editors. *Proceedings of the Twelfth IEEE Workshop on Neural Networks for Signal Processing (NNSP)*. IEEE Press, 2002.

[273] H. Bourlard and S. Bengio. Hidden markov models and other finite state automata for sequence processing. In Michael A. Arbib, editor, *The Handbook of Brain Theory and Neural Networks, Second Edition*. The MIT Press, 2002.

[274] S. Bengio, Y. Bengio, J. Cloutier, and J. Gecsei. On the optimization of a synaptic learning rule. In D. S. Levine and W. R. Elsberry, editors, *Optimality in Biological and Artificial Networks?*, pages 265–287. Lawrence Erlbaum Associates, 1997.

[275] S. Bengio. *Optimisation d'une rÃ¨gle d'apprentissage pour rÃ©seaux de neurones artificiels*. PhD thesis, DÃ©partement d'Informatique et Recherche OpÃ©rationnelle. UniversitÃ© de MontrÃ©al, 1993.

[276] S. Bengio. IntÃ©gration des systÃ¨mes tutoriels traditionnels et des systÃ¨mes tutoriels intelligents. Master's thesis, DÃ©partement d'Informatique et de Recherche OpÃ©rationnelle, UniversitÃ© de MontrÃ©al, 1989.

Selection of Technical Reports

[277] D. Grangier and S. Bengio. A discriminative approach for the retrieval of images from text queries. Technical Report IDIAP-RR 06-15, IDIAP, 2006.

[278] M. Keller and S. Bengio. A multitask learning approach to document representation using unlabeled data. Technical Report IDIAP-RR 06-44, IDIAP, 2006.

[279] S. Bengio. Joint training of multi-stream HMMs. Technical Report IDIAP-RR 05-22, IDIAP, 2005.

[280] D. Grangier and S. Bengio. A discriminative decoder for the recognition of phoneme sequences. Technical Report IDIAP-RR 05-67, IDIAP, 2005.

[281] J. MariÃ©thoz and S. Bengio. Can a professional imitator fool a gmm-based speaker verification system? Technical Report IDIAP-RR 05-61, IDIAP, 2005.

[282] A. Pozdnoukhov and S. Bengio. A kernel classifier for distributions. Technical Report IDIAP-RR 05-32, IDIAP, 2005.

[283] D. Zhang, D. Gatica-Perez, D. Roy, and S. Bengio. Modeling interactions from email communication. Technical Report IDIAP-RR 05-51, IDIAP, 2005.

[284] S. Chiappa and S. Bengio. Sequence classification with input-output hidden markov models. Technical Report IDIAP-RR 04-13, IDIAP, 2004.

[285] C. Dimitrakakis and S. Bengio. Estimates of parameter distributions for optimal action selection. Technical Report IDIAP-RR 04-72, IDIAP, 2004.

[286] M. Keller, J. MariÃ©thoz, and S. Bengio. Significance tests for bizarre measures in 2-class classification tasks. Technical Report IDIAP-RR 04-34, IDIAP, 2004.

[287] C. Sanderson and S. Bengio. Statistical transformation techniques for face verification using faces rotated in depth. Technical Report IDIAP-RR 04-04, IDIAP, 2004.

[288] R. Collobert and S. Bengio. A new margin-based criterion for efficient gradient descent. Technical Report IDIAP-RR 03-16, IDIAP, 2003.

[289] M. Keller and S. Bengio. Textual data representation. Technical Report IDIAP-RR 03-49, IDIAP, 2003.

[290] J. MariÃ©thoz and S. Bengio. An alternative to silence removal for text-independent speaker verification. Technical Report IDIAP-RR 03-51, IDIAP, 2003.

[291] I. McCowan, D. Gatica-Perez, and S. Bengio. Meeting data collection specifications. Communication Report IDIAP-COM 03-10, IDIAP, 2003.

[292] N. Poh and S. Bengio. Variance reduction techniques in biometric authentication. Technical Report IDIAP-RR 03-17, IDIAP, 2003.

[293] A. Pozdnoukhov and S. Bengio. From samples to objects in kernel methods. Technical Report IDIAP-RR 03-29, IDIAP, 2003.

[294] Q. Le and S. Bengio. Hybrid generative-discriminative models for speech and speaker recognition. Technical Report IDIAP-RR 02-06, IDIAP, 2002.

[295] F. PorÃ©e, J. MariÃ©thoz, S. Bengio, and F. Bimbot. The BANCA database and experimental protocol for speaker verification. Technical Report IDIAP-RR 02-13, IDIAP, 2002.

[296] A. Vinciarelli and S. Bengio. Transforming the feature vectors to improve HMM based cursive word recognition systems. Technical Report IDIAP-RR 02-32, IDIAP, 2002.

[297] S. Bengio and J. MariÃ©thoz. Comparison of client model adaptation schemes. Technical Report IDIAP-RR 01-25, IDIAP, 2001.

[298] S. Bengio, J. MariÃ©thoz, and S. Marcel. Evaluation of biometric technology on XM2VTS. Technical Report IDIAP-RR 01-21, IDIAP, 2001.

[299] K. Weber, S. Bengio, and H. Bourlard. A pragmatic view of the application of HMM2 for ASR. Technical Report IDIAP-RR 01-23, IDIAP, 2001.

[300] S. Bengio, H. Bourlard, and K. Weber. An EM algorithm for HMMs with emission distributions represented by HMMs. Technical Report IDIAP-RR 00-11, IDIAP, Martigny, Switzerland, 2000.

[301] R. Collobert and S. Bengio. On the convergence of SVMTorch, an algorithm for large-scale regression problems. Technical Report IDIAP-RR 00-24, IDIAP, Martigny, Switzerland, 2000.

[302] R. Collobert and S. Bengio. Support vector machines for large-scale regression problems. Technical Report IDIAP-RR 00-17, IDIAP, Martigny, Switzerland, 2000.

[303] A. Gravey, S. Bengio, D. Collobert, and F. Clerot. Utilisation de techniques de prÃ©diction neuromimÃ©tiques pour la nÃ©gotiation dynamique des paramÃ¨tres de contrat de trafic dans un rÃ©seau ATM. Technical Report NT/LAA/EIA/132, France TÃ©lÃ©com CNET, Lannion, France, 1996.

[304] J.-Y. Potvin and S. Bengio. A genetic based heuristic for the vehicle routing problem with time windows. Technical Report CRT-953, Centre de Recherche sur les Transports, UniversitÃ© de MontrÃ©al, 1994.

[305] Y. Bengio and S. Bengio. Learning a synaptic learning rule. Technical Report 751, DÃ©partement d'Informatique et de Recherche OpÃ©rationnelle, UniversitÃ© de MontrÃ©al, Montreal (QC) Canada, 1990.

[306] S. Bengio and C. Frasson. Utilisation d'EAO dans des systÃ¨mes d'EIAO. Technical Report 651, DÃ©partement d'Informatique et de Recherche OpÃ©rationnelle, UniversitÃ© de MontrÃ©al, Montreal (QC) Canada, 1988.

[307] S. Bengio, G. Brassard, Y. Desmedt, C. Goutier, and J.-J. Quisquater. Aspects and importance of secure implementations of identification systems. Technical Report Manuscript M209, Philips Research Laboratory, Brussel, Belgium, 1987.