Yoshua Bengio - C.V.

Département d'informatique
et recherche opérationnelle (DIRO),
Université de Montréal

Phone : 514.343.6804
yoshua.bengio@mila.quebec
https://yoshua-bengio.org/

A.M. Turing Award,
Fellow of the Royal Society of London,
Fellow of the Royal Society of Canada,
Canada CIFAR AI Chair

Titles and Distinctions

• **Full Professor**, Université de Montréal, department of computer science and operations research, since 2002. Recruited as assistant prof. in 1993. Previously associate professor (1997-2002) and assistant professor (1993-1997).

• **Canada Research Chair** on *Statistical Learning Algorithms* 2001-2019. **Recipient of a Canada CIFAR AI Chair (CCAI)** (2018-2023, 1.25M$).

• **Recipient** of the 2018 A.M. Turing Award, of the 2019 Killam Prize, of the 2017 Marie-Victorin Québec Prize (highest distinction in the sciences, for the province of Québec), Prix d'excellence du FRQNT 2019, Medal of the 50th Anniversary of the Ministry of International Relations and Francophonie (2018), 2019 IEEE CIS (Computational Intelligence Society) Neural Networks Pioneer Award, ACFAS Urgel-Archambault 2009 prize (covering physics, mathematics, computer science, and engineering). Radio-Canada's Scientist of the Year (2017), Lifetime Achievement Award 2018 from the Canadian AI Association.

• **Founder and scientific director of Mila**, the Quebec Artificial Intelligence Institute (formerly the Montreal Institute for Learning Algorithms), which brings together the researchers of U. Montréal, HEC, Polytechnique Montreal and McGill in an independent non-profit organization now counting over 500 researchers, including 80 faculty. Mila is one of the three federally-funded centres of excellence in AI research and innovation in Canada. It is the largest academic center for deep learning research in the world, yielding pioneering papers of the area, starting with the introduction of deep learning in 2006, curriculum learning in 2009, showing the power of ReLUs for deeper nets in 2011, and the breakthroughs brought by GANs, neural machine translation, and attention mechanisms in 2014-2015.

• **Scientific Director of IVADO**, the Data Valorization Institute, and leading applicant in the 93.6M$ CFREF grant (2016-2023) on *Data Serving Canadians : Deep Learning and*

Optimization for the Knowledge Revolution, largest grant ever received at U. Montreal.

• **Fellow of the Royal Society of London** (2020), **Fellow of the Royal Society of Canada** (2017) and named **Officer of the Order of Canada** in 2017.

• **Co-Chair of the Advisory Council on Artificial Intelligence of Canada** since 2018. **Co-Chair of the Global Partnership on Artificial Intelligence (GPAI)** Working Group on Responsible AI since 2020.

• **Senior Fellow**, CIFAR (Canadian Institute For Advanced Research), since 2004. **Co-director** of the CIFAR LMB (Learning in Machines in Brains) program (previously called NCAP) since 2014. This is the program which funded the initial breakthroughs in deep learning and was originally led by Geoff Hinton. **NSERC Industrial Research Chair** 2005-2015. **Canada CIFAR AI Chair** since 2019.

• Member of the advisory board of the **Neural Information Processing Systems (NeurIPS) Foundation**, since 2010, and member of the board of the **International Conference on Learning Representations** (ICLR), since 2013.

• **Action Editor**, *Journal of Machine Learning Research* (JMLR), *Neural Computation*, *Foundations and Trends in Machine Learning*, and *Computational Intelligence*. Member of the 2012 editor-in-chief nominating committee for JMLR. Previously **Associate Editor**, *Machine Learning*, *IEEE Trans. on Neural Networks*. *La Recherche* chose his work on neural networks local minima as one of the 10 highlighted discoveries of 2015.

• Member of the board of the Centre de Recherches Mathématiques, UdeM, 1999-2009. Head of the CRM-MILA Lab. Member of the Awards Committee of the Canadian Association for Computer Science (2012-2013). Member of the NIPS'2012 and ICML'2016 committees for best paper awards, and NIPS committees for choosing the next program chairs 2013-2018.

• *Program Co-Chair*, NIPS'2008, *General Chair*, NIPS'2009. Note that *Advances in Neural Information Processing Systems* (NeurIPS) is a very high level conference, the most important in my field (> 1000 submissions), with reviewing and acceptance criteria comparing favorably to best journals (acceptance rate between 20% and 25%). I am one of the most prolific NeurIPS authors (56 papers).

• Member of grant selection committees for Quebec's FQRNT (1999-2000), Canada's NSERC (2000-2003, 2006-2007), and CACS/AIC Awards Committee 2012-2015.

• Co-founder of multiple start-ups, including Element AI (2016), which raised a record-breaking 135M$ for its series A. Leading the effort to connect the Mila with the AI entrepreneurial ecosystem and make Montreal the AI hub of the world, bringing to

Montreal AI research labs of Microsoft, Google, Facebook, DeepMind and Samsung.

• Over 380 invited talks, at places such as the NIPS'2015 tutorial, many NIPS workshops, ICML workshops, Stanford, Berkeley, Oxford, Cambridge, Google Research, UCL, CMU, Cornell, Deep Mind, Facebook AI Research, Microsoft Research, Open AI, the Fields Institute, MIT, University College London, New York University, Johns Hopkins U., U. Toronto, IBM Research, Intel Research, Yahoo Research, Qualcomm, Samsung, Nuance, Twitter, the Gatsby Unit, the ICONIP conference, DARPA, ETH, IDIAP, summer schools. Tutorials at ACL 2012, AAAI 2013, IPAM, SSTiC 2013, NIPS'2015, ECCV'2016.

• Creation of the ICLR (Int. Conf. on Learning Representations) in 2013, general chair for ICLR'2013, ICLR'2014, ICLR'2015, ICLR'2016, ICLR'2017, organization or co-organization of more than 21 workshops in career, including the deep or representation learning workshops or symposia at NIPS'2007, ICML'2009, NIPS'2010, 2011, 2012, 2013, 2014 and 2015, at ICML'2012, 2013 and 2015. Co-organizing the CIFAR-CRM Deep Learning summer schools (2015, 2016, 2017).

Major Scientific Impact

In September 2019, there were nearly 206 000 citations to scientific publications authored by Yoshua Bengio found by *Google Scholar* (over 395 000 by 04/2021), with an H-index of 146 (183 by 04/2021), with over 60 000 citations in 2018 alone. As recognized in particular by the Turing Award, his main contributions co-created the field of deep learning, in areas such as recurrent nets, methods enabling deep learning to succeed, theoretical understanding of deep learning, the development of novel architectures based on attention and making it possible for neural nets to process sets rather than just vectors and sequences, and the development of deep generative models such as the generative adversarial networks. More recently, Y.B. turned his attention to the agent perspective for representation learning (and thus to deep reinforcement learning) and deep learning architectures for reasoning, causality and systematic generalization, and has taken part in national and global discussions (and documents) about the socially responsible development of AI, as well as contributing to the research on AI for social good applications, e.g. in healthcare, the environment and education. Major contributions are highlighted below.

• 1989-1998 Convolutional and recurrent networks combined with probabilistic alignment (HMMs) to model sequences, as the main contribution of my PhD thesis (1991), NIPS'1988, NIPS'1989, Eurospeech'1991, PAMI'1991, IEEE Trans. Neural Nets 1992. These architectures were first applied to **speech recognition** in my PhD (and rediscovered after 2010) and then with Yann LeCun et al to **handwriting recognition and document analysis** (most cited paper is 'Gradient-based learning applied to document recognition', 1998, with over 19 000 citations).

• 1991-1995 **Learning to learn** papers with Samy Bengio, starting with IJCNN 1991, "Learning a synaptic learning rule". The idea of learning to learn (in particular by back-propagating through the whole process) has now become very popular (now called meta-learning) but we lacked the necessary computing power in the early 90's.

• 1993-1995 Uncovering the **fundamental difficulty of learning in recurrent nets** and

other machine learning models of temporal dependencies, associated with vanishing and exploding gradients : ICNN'1993, NIPS'1993, NIPS'1994, IEEE Transactions on Neural Nets 1994, NIPS'1995. These papers (in particular the negative result) have had a major impact (turning the field of recurrent nets upside down) and motivated later papers on architectures to help learning long-term dependencies and deal with vanishing or exploding gradients. An important but subtle contribution of the IEEE Transactions 1994 paper is to show that the condition required to store bits of information reliably over time also gives rise to vanishing gradients, using dynamical systems theory. The NIPS'1995 paper introduced the use of a hierarchy of time scales to combat the vanishing gradients issue.

• 1999-2014 Understanding how **distributed representations** can bypass the **curse of dimensionality** by providing generalization to an exponentially large set of regions from those comparatively few occupied by training examples. This series of papers also highlights how methods based on local generalization, like nearest-neighbor and Gaussian kernel SVMs lack this kind of generalization ability. The NIPS'1999 introduced for the first time auto-regressive neural networks for density estimation (the ancestor of the NADE and PixelRNN/PixelCNN models). The NIPS'2004, NIPS'2005 and NIPS'2011 papers on this subject show how neural nets can learn a local metric which can bring the power of generalization of distributed representations to kernel methods and manifold learning methods. Another NIPS'2005 paper shows the fundamental limitations of kernel methods due to a generalization of the curse of dimensionality (the curse of highly variable functions, which have many ups and downs). Finally, the ICLR'2014 paper shows in the case of piecewise-linear networks (like those with ReLUs) that the regions (linear pieces) distinguished by a one-hidden layer network is exponential in the number of neurons (whereas the number of parameters is quadratic in the number of neurons, and a local kernel method would require an exponential number of examples to capture the same kind of function).

• 2000-2008 **Word embeddings from neural networks and neural language models**. The NIPS'2000 paper introduces for the first time the learning of word embeddings as part of a neural network which models language data. The JMLR'2003 journal version expands this (these two papers together get around 3000 citations) and also introduces the idea of **asynchronous SGD** for distributed training of neural nets. Word embeddings have become one of the most common fixtures of deep learning when it comes to language data and this has basically created a new sub-field in the area of computational linguistics. I also introduced the use of importance sampling (AISTATS'2003, IEEE Trans. on Neural Nets, 2008) as well as of a probabilistic hierarchy (AISTATS 2005) to speed-up computations and face larger vocabularies.

• 2006-2014 Showing the **theoretical advantage of depth** for generalization. The NIPS'2006 oral shows experimentally the advantage of depth and is one of the most cited papers in the field (over 2600 citations). The NIPS'2011 paper shows how deeper sum-product networks can represent functions which would otherwise require an exponentially larger model if the network is shallow. Finally, the NIPS'2014 paper on the number of linear regions of deep neural networks generalizes the ICLR'2014 paper mentionned above, showing that the number of linear pieces produced by a piecewise linear network grows exponentially in both width of layers and number of layers, i.e., depth, making the functions represented by such networks generally impossible to capture efficiently

with kernel methods (short of using a trained neural net as the kernel).
• 2006-2014 **Unsupervised deep learning** based on auto-encoders (with the special case of GANs as decoder-only models, see below). The NIPS'2006 paper introduced greedy layer-wise pre-training, both the in the supervised case and the unsupervised case with auto-encoders. The ICML'2008 paper introduced **denoising auto-encoders** and the NIPS'2013, ICML'2014 and JMLR'2014 papers cast their their theory and generalize them as proper probabilistic models, at the same time introducing alternatives to maximum likelihood as training principles.
• 2014 Dispelling the **local-minima myth** regarding the optimization of neural networks, with the NIPS'2014 paper on saddle points, showing that it is the large number of parameters which makes it very unlikely that bad local minima exist.
• 2014 Introducing **Generative Adversarial Networks (GANs)** at NIPS'2014, which innovates in many ways to train deep generative models, outside of the maximum likelihood framework and even outside of the classical framework of having a single objective function (instead entering into the territory of multiple models trained in a game-theoretical way, each with their objective). One of the hottest research areas in deep learning, as of this writing, with almost 2000 citations mostly from papers which introduce variants of GANs, which have been producing impressively realistic synthetic images one would not imagine computers being able to generate just a few years ago.
• 2014-2016 Introducing **content-based soft attention** and the breakthrough it brought to **neural machine translation**, mostly with Kyunghyun Cho and Dima Bahdanau. We first introduced the encoder-decoder (now called sequence-to-sequence) architecture (EMNLP'2014) and then achieved a big jump in BLEU scores with content-based soft attention (ICLR'2015). These ingredients are now the basis of most commercial machine translation systems. Another whole sub-field has been created using these techniques.

Academic Studies and Diplomas

1992–1993	**Post-doctoral Fellow, Learning and vision algorithms,** *Larry Jackel and Yann LeCun's group,* AT&T Bell Laboratories, New-Jersey.
1991–1992	**Post-doctoral Fellow,** *NSERC scholarship,* **Statistical learning / sequential data,** *Michael I. Jordan's group,* Brain and Cognitive Sciences Dept., MIT, Massachusetts.
1988–1991	**Ph.D. in computer science,** *NSERC scholarship,* **Neural Networks and Markovian Models,** Computer Science Dept., McGill University.
1986–1988	**M.Sc. in computer science** , CGPA 4.0/4.0, *Speech recognition with statistical methods,* Computer Science Dept., McGill University.
1982–1986	**B.Eng. in computer engineering,** *Honours class, CGPA 3.82/4.0* Electrical Engineering Dept., McGill University.

Graduate Students & Postdocs

Current :

Postdoc : Jason Hartford, Jie Fu, Alejandro Hernandez-Garcia, Chen Sun, Damjan Kalajdzievski, Travis Lacroix, Dianbo Liu, Sasha Luccioni, Mirco Ravanelli.

PhD : Taesup Kim, Tong Che, Sherjil Ozair, Akram Erraqabi, Valentin Thomas, William Fedus, Giancarlo Kerg, Salem Lahlou, Kundan Kumar, Alex Lamb, Anirudh Goyal, Rémi Le Priol, Tristan Deleu, Kanika Madan, Camille Rochefort-Boulanger, Ruixiang Zhang, Victor Schmidt, Paul Bertin, Prateek Gupta, Salah Al Rifai, Rosemary Nan Ke, Mohammad Pezeshki, Nasim Rahaman, Zhen Liu, Brayden Neal, Rim Assouel, Mélisande Teng, Harry (Mingde) Zhao, Maksym Korbalyov.

MSc : Leo Feng, Sarthak Mittal, Tianyu Zhang.

Former (graduated) :

Postdoc : Jonathan Binas (2021), Ghouthi Boukli Hacene (2021), Qicheng Lao (2021), Min Lin (2020), Simon Verret (2020), Maximilian Puelma Touzel (2020), Joseph Paul Cohen (2020), Jason Jo (2020), Margaux Luck (2020), Karthik Mukkavilli (2020), Kris Sankaran (2020), Devansh Arpit (2019), Devon Hjelm (2018), Nicolas Ballas (2017), Philemon Brakel (2017), Adriana Romero Soriano (2017), Sungjin Ahn (2016), Asja Fischer (2016), Jorg Bornschein (2015), Kyung-Hyun Cho (2015), Jyri Kivinen (2014), Heng Luo (2013), Antoine Bordes (2011), Aaron Courville (2011), Michael Mendel (2010), Joseph Turian (2010), Jerome Louradour (2008), Marina Sokolova (2007), Pierre-Jean L'Heureux (2006), Christopher Kermorvant (2005), Xiangdong Wang (2003), Gilles Caporossi (2002), Takafumi Kanamori (2001), Ichiro Takeuchi (2001), Stephen Langdell (2000), Claude Nadeau (2000), Holger Schwenk (1997), Samy Bengio (1996).

PhD : Sandeep Subramanian (2021), Antoine Prouvost (2021), Tristan Sylvain (2021), Benjamin Scellier (2021), Jessica Thompson (2020), Dzmitry Bahdanau (2020), Chinnadhurai Sankar (2020), Dmitriy Serdyuk (2020), Guillaume Alain (2019), Sarath Chandar (2019), Yaroslav Ganin (2019), Zhouhan Lin (2019), Bart Merrienboer (2019), Julian Vlad Serban (2019), Saizheng Zhang (2019), Junyoung Chung (2018), Laurent Dinh (2018), Vincent Dumoulin (2018), Caglar Gulcehre (2018), David Warde-Farley (2017), Li Yao (2017), Mehdi Mirza (2017), Yann Dauphin (2015), Xavier Glorot (2015), Razvan Pascanu (2014), Ian Goodfellow (2014), Guillaume Desjardins (2014), Nicolas Boulanger-Lewandoski (2013), Philippe Hamel (2012), Olivier Delalleau (2012), James Bergstra (2011), Dumitru Erhan (2011), François Rivest (2010), Nicolas Chapados (2009), Hugo Larochelle (2009), Nicolas Le Roux (2008), Julie Carreau (2008), Narjes Boufaden (2005), Pascal Vincent (2003), Charles Dugas (2003), Joumana Ghosn (2002), Steven Pigeon (2001), François Gingras (1999).

MSc : Rim Assouel (2020), Barghav Kanuparthi (2020), Stephanie Larocque (2020),

Samuel Lavoie-Marchildon (2020), Rahul Mihir Patel (2020), Rithesh Kumar (2019), Shagun Sodhani (2019), Olexa Bilaniuk (2018), Francis Dutil (2018), Dong-Hyun Lee (2018), Kelvin Xu (2017), Soroush Mehri (2016), Samira Shabanian (2016), Jose Rodriguez Sotelo (2016), Kyle Kastner (2016), David Krueger (2016), Matthieu Courbariaux (2015), Pierre Luc Carrier (2014), Eric Thibodeau-Laufer (2014), Nicholas Leonard (2014), Valentin Bisson (2012), François Savard (2011), Olivier Breuleux (2010), Guillaume Desjardins (2009), Pierre-Antoine Manzagol (2007), Dumitru Erhan (2006), Marie Ouimet (2004), Christian Dorion (2004), Maryse Boisvert (2004), Fréderic Morin (2004), Francis Piérault (2003), Jean-François Paiement (2003), Jean-Sébastien Senecal (2003), Lynian Meng (2002), Nicolas Chapados (2000), Vincent-Philippe Lauzon (1999), Simon Latendresse (1999), Julien Desaulnier (1998).

Main Grants

Current

- Fonds de la recherche en santé du Québec (FRQS), Learning causal models in single cell dynamics, $100k, 2020-2025
- NSERC Discovery grant, $89k per year $\times$ 5 yrs, 2019-2024
- IBM - USA, Mila AIHN Open collaboration open source, USD$1M, 2021-2023
- Programme de coopération climatique internationale (PCCI), $840k, 2021-2023
- Canada CIFAR AI Chair, research grant, $175k per year $\times$ 5 yrs, 2018-2023
- CFREF grant (Data for Canadians), $93.6M, 2016-2023
- NSERC Strategic Network grant, $5.5M over 5 years, 2017-2022
- Imagia Collaborative R&D grant in healthcare, $300k over 6 years, 2016-2022
- Gates Foundation, RECOVER - Gates I, USD$170k, 2020-2021
- Scale AI, Deep RL for Discovering Novel COVID-19 Antivirals, $125k, 2020-2021
- IVADO, Deep RL for Discovering Novel COVID-19 Antivirals, $100k, 2020-2021
- Facebook, unrestricted equipment gift, $1.5M, 2017
- Microsoft, unrestricted gift, $1.2M per year $\times$ 5 yrs, 2016-2021
- Samsung GRP DL grant, US$550k per year $\times$ 3 yrs, 2017-2020
- Samsung GRP NPP grant, $100k per year $\times$ 3 yrs, 2015-2018
- Nuance Foundation grants (2), 2 $\times$ $200k per year $\times$ 4 yrs, 2014-2018

Previous (past 5 years) in addition to above

- Canada Research Chair, $200k per year, 2006-2020
- CFI Cyberinfrastructure grant, $5M, 2016-2019
- Google focused research award, USD$250k per year, 2016-2019
- NSERC discovery grant, $76k per year $\times$ 5 yrs, 2014-2019
- NSERC CRD grants (with IBM as partner), $200k per year, 2016-2018
- NSERC + IBM collaborative R&D grant, $800k over 3 years, 2015-2018
- Panasonic, unrestricted gift, USD$200k in 2017, USD$300k in 2018

— Samsung GRP DL grant, $500k per year × 2 yrs, 2014-2016
— NSERC strategic grants (2), $240k +$220k per year × 3 yrs, 2013-2016
— NSERC-Ubisoft CRD grants, $50k+$80k per year, 2011-2016
— NSERC-Ubisoft industrial chair, $350k per year × 5 yrs, 2011-2016
— Panasonic research sponsorship, $250k, 2016
— NSERC equipment grant, $135k, 2016

Partial List of Co-Authors

Yann LeCun, Geoff Hinton, Aaron Courville, Pascal Vincent, Vladimir Vapnik, Leon Bottou, Hugo Larochelle, Ronan Collobert, Ian Goodfellow, Antoine Bordes, Nicolas Le Roux, Samy Bengio, James Bergstra, Yves Grandvalet, Xavier Glorot, Jason Weston, Douglas Eck, Marco Gori, Juergen Schmidhuber, Dumitru Erhan, Olivier Chapelle, Lise Getoor, Thomas Breuel, Joseph Turian, Patrice Marcotte, Balazs Kegl, Tomas Mikolov, David Warde-Farley, Guido Montufar, Gal Chechik, Andrew Fitzgibbon, Patrick Haffner, Razvan Pascanu, Guillaume Desjardins, Patrice Simard, Salah Rifai, Pascal Lamblin, Kyunghyun Cho, Heng Luo, Yann Dauphin, Jean-Luc Gauvain, Renato De Mori, Paolo Frasconi, Caglar Gulcehre, Dzmitry Bahdanau, Jason Yosinski, Frederic Bastien, Jan Chorowski, Jorg Bornschein, Gregoire Mesnil, Nicolas Boulanger-Lewandowski, Junyoung Chung, Li Yao, Kelvin Xu, Alessandro Sordoni, Sherjil Ozair, Richard Zemel, Sepp Hochreiter, Saizheng Zhang, Dmitriy Serkyuk, Vincent Dumoulin, Chris Pal, Joelle Pineau, Jamie Kiros, Asja Fischer, Jeff Clune, Li Deng, Bing Xu, Laurent Dinh, Takeuchi Ichiro, Patrice Marcotte, Felix Hill, Heng Luo, Nicholas Leonard, Stephan Gouws.

Yoshua Bengio's Publication List, July 2021

Refereed Journal Publications

[1] Gesa HARTWIGSEN, Yoshua BENGIO et Danilo BZDOK. "How does hemispheric specialization contribute to human-defining cognition?" In : *Neuron* (2021).

[2] Axel LABORIEUX, Maxence ERNOULT, Benjamin SCELLIER, Yoshua BENGIO, Julie GROLLIER et Damien QUERLIOZ. "Scaling equilibrium propagation to deep convnets by drastically reducing its gradient estimator bias". In : *Frontiers in neuroscience* 15 (2021), p. 129.

[3] Alexandra LUCCIONI, Victor SCHMIDT, Vahe VARDANYAN et Yoshua BENGIO. "Using Artificial Intelligence to Visualize the Impacts of Climate Change". In : *IEEE Computer Graphics and Applications* 41.1 (2021), p. 8–14.

[4] Yoshua BENGIO, Andrea LODI et Antoine PROUVOST. "Machine learning for combinatorial optimization: A methodological tour d'horizon". In : *European Journal of Operational Research* 290.2 (2021), p. 405–421.

[5] Qicheng LAO, Xiang JIANG, Mohammad HAVAEI et Yoshua BENGIO. "A Two-Stream Continual Learning System With Variational Domain-Agnostic Feature Replay." In : *IEEE Transactions on Neural Networks and Learning Systems* (2021).

[6] Shagun SODHANI, Sarath CHANDAR et Yoshua BENGIO. “Toward Training Recurrent Neural Networks for Lifelong Learning”. In : *Neural computation* 32.1 (2020), p. 1–35.

[7] Iulian Vlad SERBAN, Chinnadhurai SANKAR, Michael PIEPER, Joëlle PINEAU et Yoshua BENGIO. “The bottleneck simulator: A model-based deep reinforcement learning approach”. In : *Journal of Artificial Intelligence Research* 69 (2020), p. 571–612.

[8] Joseph Paul COHEN, Lan DAO, Karsten ROTH, Paul MORRISON, Yoshua BENGIO, Almas F ABBASI, Beiyi SHEN, Hoshmand Kochi MAHSA, Marzyeh GHASSEMI, Haifang LI et Tim DUONG. “Predicting COVID-19 Pneumonia Severity on Chest X-ray With Deep Learning”. In : *Cureus* 12.7 (2020).

[9] Assya TROFIMOV, Joseph Paul COHEN, Yoshua BENGIO, Claude PERREAULT et Sébastien LEMIEUX. “Factorized embeddings learns rich and biologically meaningful embedding spaces using factorized tensor decomposition”. In : *Bioinformatics* 36.Supplement_1 (2020), p. i417–i426.

[10] Yoshua BENGIO, Daphne IPPOLITO, Richard JANDA, Max JARVIE, Benjamin PRUD’HOMME, Jean-François ROUSSEAU, Abhinav SHARMA et Yun William YU. “Inherent privacy limitations of decentralized contact tracing apps”. In : *Journal of the American Medical Informatics Association* (2020).

[11] Yoshua BENGIO, Richard JANDA, Yun William YU, Daphne IPPOLITO, Max JARVIE, Dan PILAT, Brooke STRUCK, Sekoul KRASTEV et Abhinav SHARMA. “The need for privacy with public digital contact tracing during the COVID-19 pandemic”. In : *The Lancet Digital Health* (2020).

[12] Alexandra LUCCIONI et Yoshua BENGIO. “On the Morality of Artificial Intelligence [Commentary]”. In : *IEEE Technol. Soc. Mag.* (2020), p. 16–25.

[13] Konrad WAGSTYL, Stéphanie LAROCQUE, Guillem CUCURULL, Claude LEPAGE, Joseph Paul COHEN, Sebastian BLUDAU, Nicola PALOMERO-GALLAGHER, Lindsay B LEWIS, Thomas FUNCK, Hannah SPITZER et al. “BigBrain 3D atlas of cortical layers: Cortical and laminar thickness gradients diverge in sensory and motor cortices”. In : *PLoS biology* 18.4 (2020), e3000678.

[14] Sharon ZHOU, Alexandra LUCCIONI, Gautier COSNE, Michael S BERNSTEIN et Yoshua BENGIO. “Establishing an evaluation metric to quantify climate change image realism”. In : *Machine Learning: Science and Technology* 1.2 (2020), p. 025005.

[15] Benjamin SCELLIER et Yoshua BENGIO. “Equivalence of equilibrium propagation and recurrent backpropagation”. In : *Neural computation* 31.2 (2019), p. 312–329.

[16] Blake A RICHARDS, Timothy P LILLICRAP, Philippe BEAUDOIN, Yoshua BENGIO, Rafal BOGACZ, Amelia CHRISTENSEN, Claudia CLOPATH, Rui Ponte COSTA, Archy de BERKER, Surya GANGULI et al. “A deep learning framework for neuroscience”. In : *Nature neuroscience* 22.11 (2019), p. 1761–1770.

[17] Kenji KAWAGUCHI et Yoshua BENGIO. “Depth with nonlinearity creates no bad local minima in ResNets”. In : *Neural Networks* 118 (2019), p. 167–174.

[18] Li JING, Caglar GULCEHRE, John PEURIFOY, Yichen SHEN, Max TEGMARK, Marin SOLJACIC et Yoshua BENGIO. “Gated orthogonal recurrent units: On learning to forget”. In : *Neural computation* 31.4 (2019), p. 765–783.

[19] Vincent DUMOULIN, Ethan PEREZ, Nathan SCHUCHER, Florian STRUB, Harm de VRIES, Aaron COURVILLE et Yoshua BENGIO. “Feature-wise transformations”. In : *Distill* 3.7 (2018), e11.

[20] Caglar GÜLÇEHRE, Sarath CHANDAR, Kyunghyun CHO et Yoshua BENGIO. “Dynamic Neural Turing Machine with Continuous and Discrete Addressing Schemes”. In : *Neural Computation* 30.4 (mar. 2018). Sous la dir. de Neural COMPUTATION, p. 857–884.

[21] Georgy DEREVYANKO, Sergei GRUDININ, Yoshua BENGIO et Guillaume LAMOUREUX. "Deep convolutional networks for quality assessment of protein folds". In : *Bioinformatics* 34.23 (2018), p. 4046–4053.

[22] Mirco RAVANELLI, Philemon BRAKEL, Maurizio OMOLOGO et Yoshua BENGIO. "Light Gated Recurrent Units for Speech Recognition". In : *IEEE Transactions on Emerging Topics in Computational Intelligence* 2.2 (2018), 92–102, outstanding paper award.

[23] Heeyoul CHOI, Kyunghyun CHO et Yoshua BENGIO. "Fine-grained attention mechanism for neural machine translation". In : *Neurocomputing* 284 (2018), p. 171–176.

[24] Xu-Yao ZHANG, Fei YIN, Yan-Ming ZHANG, Cheng-Lin LIU et Yoshua BENGIO. "Drawing and recognizing chinese characters with recurrent neural network". In : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40.4 (2018), p. 849–862.

[25] Michal DROZDZAL, Gabriel CHARTRAND, Eugene VORONTSOV, Mahsa SHAKERI, Lisa Di JORIO, An TANG, Adriana ROMERO, Yoshua BENGIO, Chris PAL et Samuel KADOURY. "Learning normalized inputs for iterative estimation in medical image segmentation". In : *Medical image analysis* 44 (2018), p. 1–13.

[26] Itay HUBARA, Matthieu COURBARIAUX, Daniel SOUDRY, Ran EL-YANIV et Yoshua BENGIO. "Quantized neural networks: Training neural networks with low precision weights and activations". In : *The Journal of Machine Learning Research* 18.1 (2017), p. 6869–6898.

[27] Phil DE LUNA, Jennifer WEI, Yoshua BENGIO, Alán ASPURU-GUZIK et Edward SARGENT. *Use machine learning to find energy materials*. 2017.

[28] Felix HILL, Kyunghyun CHO, Sébastien JEAN et Yoshua BENGIO. "The representational geometry of word meanings acquired by neural machine translation models". In : *Machine Translation* 31 (2017), p. 1–16.

[29] Mohammad HAVAEI, Axel DAVY, David WARDE-FARLEY, Antoine BIARD, Aaron COURVILLE, Yoshua BENGIO, Christopher PAL, Pierre-Marc JODOIN et Hugo LAROCHELLE. "Brain tumor segmentation with Deep Neural Networks". In : *Medical Image Analysis* 35 (2017), p. 18–31.

[30] Yoshua BENGIO, Thomas MESNARD, Asja FISCHER, Saizheng ZHANG et Yuhuai WU. "STDP-compatible approximation of back-propagation in an energy-based model". In : *Neural Computation* 29.3 (2017), p. 555–577.

[31] Caglar GÜLÇEHRE, Orhan FIRAT, Kelvin XU, Kyunghyun CHO et Yoshua BENGIO. "On integrating a language model into neural machine translation". In : *Computer Speech & Language* 45 (2017), p. 137–148.

[32] Xue-Yao ZHANG, Yoshua BENGIO et Cheng-Lin LIU. "Online and offline handwritten Chinese character recognition: A comprehensive study and new benchmark". In : *Pattern Recognition* 61 (2017), p. 348–360.

[33] Xu-Yao ZHANG, Guo-Sen XIE, Cheng-Lin LIU et Yoshua BENGIO. "End-to-End Online Writer Identification With Recurrent Neural Network". In : *IEEE Transactions on Human-Machine Systems* 47.2 (2017), p. 285–292.

[34] Heeyoul CHOI, Kyunghyun CHO et Yoshua BENGIO. "Context-dependent word representation for neural machine translation". In : *Computer Speech & Language* 45 (2017), p. 149–160.

[35] Orhan FIRAT, Kyunghyun CHO, Baskaran SANKARAN, Fatos T Yarman VURAL et Yoshua BENGIO. "Multi-way, multilingual neural machine translation". In : *Computer Speech & Language* 45 (2016), p. 236–252.

[36] Yoshua BENGIO. "Springtime for AI: The Rise of Deep Learning". In : *Scientific American* (juin 2016).

[37] Guillaume ALAIN, Yoshua BENGIO, Li YAO, Jason YOSINSKI, Eric THIBODEAU-LAUFER, Saizheng ZHANG et Pascal VINCENT. "GSNs: generative stochastic networks". In : *Information and Inference* 5.2 (2016), p. 210–249.

[38] Felix HILL, Kyunghyun CHO, Anna KORHONEN et Yoshua BENGIO. "Learning to Understand Phrases by Embedding the Dictionary". In : *Transactions of the Association for Computational Linguistics* 4 (2016), p. 17–30.

[39] Caglar GÜLÇEHRE et Yoshua BENGIO. "Knowledge Matters: Importance of Prior Information for Optimization". In : *Journal of Machine Learning Research* 17.8 (2016), p. 1–32.

[40] Yann LECUN, Yoshua BENGIO et Geoffrey HINTON. "Deep Learning". In : *Nature* 521.7553 (2015), p. 436–444.

[41] Kyunghyun CHO, Aaron COURVILLE et Yoshua BENGIO. "Describing multimedia content using attention-based encoder-decoder networks". In : *Multimedia, IEEE Transactions on* 17.11 (2015), p. 1875–1886.

[42] Ian J. GOODFELLOW et al. "Challenges in Representation Learning: A report on three machine learning contests". In : *Neural Networks* 64 (2015), p. 59–63.

[43] Samira EBRAHIMI KAHOU et al. "EmoNets: Multimodal deep learning approaches for emotion recognition in video". In : *Journal on Multimodal User Interfaces* (2015), p. 1–13.

[44] Francois RIVEST, John F KALASKA et Yoshua BENGIO. "Conditioning and time representation in long short-term memory networks". In : *Biological cybernetics* 108.1 (2014), p. 23–48.

[45] Grégoire MESNIL, Yann DAUPHIN, Kaisheng YAO, Yoshua BENGIO, Li DENG, Dilek HAKKANI-TUR, Xiaodong HE, Larry HECK, Gokhan TUR, Dong YU et al. "Using recurrent neural networks for slot filling in spoken language understanding". In : *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 23.3 (2014), 530–539, IEEE SPS Best Paper Award.

[46] Guillaume ALAIN et Yoshua BENGIO. "What Regularized Auto-Encoders Learn from the Data-Generating Distribution". In : *Journal of Machine Learning Research* 15 (2014), p. 3563–3593.

[47] Aaron COURVILLE, Guillaume DESJARDINS, James BERGSTRA et Yoshua BENGIO. "The Spike-and-Slab RBM and Extensions to Discrete and Sparse Data Distributions". In : *IEEE Tr. PAMI* 36.9 (2014), p. 1874–1887.

[48] Yoshua BENGIO, Aaron COURVILLE et Pascal VINCENT. "Representation Learning: A Review and New Perspectives". In : *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)* 35.8 (2013), p. 1798–1828.

[49] Antoine BORDES, Xavier GLOROT, Jason WESTON et Yoshua BENGIO. "A Semantic Matching Energy Function for Learning with Multi-relational Data". In : *Machine Learning: Special Issue on Learning Semantics* 94.2 (2013), p. 233–259.

[50] Grégoire MESNIL, Antoine BORDES, Jason WESTON, Gal CHECHIK et Yoshua BENGIO. "Learning Semantic Representations Of Objects And Their Parts". In : *Machine Learning: Special Issue on Learning Semantics* 94.2 (2013), p. 281–301.

[51] Olivier DELALLEAU, Emile CONTAL, Eric THIBODEAU-LAUFER, Raul CHANDIAS FERRARI, Yoshua BENGIO et Frank ZHANG. "Beyond Skill Rating: Advanced Matchmaking in Ghost Recon Online". In : *IEEE Transactions on Computational Intelligence and AI in Games* 4.3 (sept. 2012), p. 167–177.

[52] James BERGSTRA et Yoshua BENGIO. "Random Search for Hyper-Parameter Optimization". In : *Journal of Machine Learning Research* 13 (fév. 2012), p. 281–305.

[53] Hugo LAROCHELLE, Michael MANDEL, Razvan PASCANU et Yoshua BENGIO. "Learning Algorithms for the Classification Restricted Boltzmann Machine". In : *Journal of Machine Learning Research* 13 (mar. 2012), p. 643–669.

[54] Yoshua BENGIO, Nicolas CHAPADOS, Olivier DELALLEAU, Hugo LAROCHELLE et Xavier SAINT-MLEUX. "Detonation Classification from Acoustic Signature with the Restricted Boltzmann Machine". In : *Computational Intelligence* 28.2 (2012), p. 261–288.

[55] Olivier BREULEUX, Yoshua BENGIO et Pascal VINCENT. "Quickly Generating Representative Samples from an RBM-Derived Process". In : *Neural Computation* 23.8 (août 2011), p. 2053–2073.

[56] James BERGSTRA, Yoshua BENGIO et Jerome LOURADOUR. "Suitability of V1 Energy Models for Object Classification". In : *Neural Computation* 23.3 (mar. 2011), p. 774–790.

[57] Michael MANDEL, Razvan PASCANU, Douglas ECK, Yoshua BENGIO, L. M. AEILLO, R. SCHIFANELLA et F. MENCZER. "Contextual tag inference". In : *ACM T. Multimedia Comp., Comm. & Appl.* 7S (oct. 2011), p. 1–32.

[58] Pascal VINCENT, Hugo LAROCHELLE, Isabelle LAJOIE, Yoshua BENGIO et Pierre-Antoine MANZAGOL. "Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion". In : *Journal of Machine Learning Research* 11 (déc. 2010), p. 3371–3408.

[59] Hugo LAROCHELLE, Yoshua BENGIO et Joseph TURIAN. "Tractable Multivariate Binary Density Estimation and the Restricted Boltzmann Forest". In : *Neural Computation* 22.9 (sept. 2010), p. 2285–2307.

[60] Nicolas LE ROUX et Yoshua BENGIO. "Deep Belief Networks are Compact Universal Approximators". In : *Neural Computation* 22.8 (août 2010), p. 2192–2207.

[61] Yoshua BENGIO, Olivier DELALLEAU et Clarence SIMARD. "Decision Trees do not Generalize to New Variations". In : *Computational Intelligence* 26.4 (nov. 2010), p. 449–467.

[62] Dumitru ERHAN, Yoshua BENGIO, Aaron COURVILLE, Pierre-Antoine MANZAGOL, Pascal VINCENT et Samy BENGIO. "Why Does Unsupervised Pre-training Help Deep Learning?" In : *Journal of Machine Learning Research* 11 (fév. 2010), p. 625–660.

[63] François RIVEST, John KALASKA et Yoshua BENGIO. "Alternative Time Representations in Dopamine Models". In : *Journal of Computational Neuroscience* 28.1 (2009), p. 107–130.

[64] Yoshua BENGIO. "Learning deep architectures for AI". In : *Foundations and Trends in Machine Learning* 2.1 (2009), p. 1–127.

[65] Charles DUGAS, Yoshua BENGIO, Francois BELISLE, Claude NADEAU et Rene GARCIA. "Incorporating Functional Knowledge in Neural Networks". In : *The Journal of Machine Learning Research* 10 (juin 2009), p. 1239–1262.

[66] Julie CARREAU et Yoshua BENGIO. "A Hybrid Pareto Mixture for Conditional Asymmetric Fat-Tailed Distribution". In : *IEEE Transactions on Neural Networks* 20.7 (2009), p. 1087–1101.

[67] Julie CARREAU et Yoshua BENGIO. "A Hybrid Pareto Model for Asymmetric Fat-Tailed Data: the univariate case". In : *Extremes* 12.1 (2009), p. 53–76.

[68] Hugo LAROCHELLE, Yoshua BENGIO, Jerome LOURADOUR et Pascal LAMBLIN. "Exploring Strategies for Training Deep Neural Networks". In : *Journal of Machine Learning Research* 10 (jan. 2009), p. 1–40.

[69] Yoshua BENGIO et Olivier DELALLEAU. "Justifying and Generalizing Contrastive Divergence". In : *Neural Computation* 21.6 (juin 2009), p. 1601–1621.

[70] Nicolas LE ROUX et Yoshua BENGIO. “Representational Power of Restricted Boltzmann Machines and Deep Belief Networks”. In : *Neural Computation* 20.6 (juin 2008), p. 1631–1649.

[71] Yoshua BENGIO et Jean-Sébastien SÉNÉCAL. “Adaptive Importance Sampling to Accelerate Training of a Neural Probabilistic Language Model”. In : *IEEE Trans. Neural Networks* 19.4 (2008), p. 713–722.

[72] Yoshua BENGIO. “Neural net language models”. In : *Scholarpedia* 3.1 (2008), p. 3881.

[73] Yoshua BENGIO. “On the Challenge of Learning Complex Functions”. In : *Computational Neuroscience: Theoretical Insights into Brain Function*. Sous la dir. de Paul CISEK, John KALASKA et Trevor DREW. T. 165. Progress in Brain Research. Elsevier, 2007, p. 521–534.

[74] Nicolas CHAPADOS et Yoshua BENGIO. “Noisy K Best-Paths for Approximate Dynamic Programming with Application to Portfolio Optimization”. In : *Journal of Computers* 2.1 (2007), p. 12–19.

[75] Yoshua BENGIO, Martin MONPERRUS et Hugo LAROCHELLE. “Nonlocal Estimation of Manifold Structure”. In : *Neural Computation* 18.10 (2006), p. 2509–2528.

[76] Dumitru ERHAN, Pierre-Jean L’HEUREUX, Shi Yi YUE et Yoshua BENGIO. “Collaborative Filtering on a Family of Biological Targets.” In : *Journal of Chemical Information and Modeling* 46.2 (2006), p. 626–635.

[77] Maria Clara ZACCARO, Hong BOON, Mookda PATTARAWARAPAN, Zebin XIA, Antoine CARON, Pierre-Jean L’HEUREUX, Yoshua BENGIO, Kevin BURGESS et H. Uri SARAGORI. “Selective Small Molecule Peptidomimetic Ligands of TrkC and TrkA Receptors Afford Discrete or Complete Neurotrophic Activities”. In : *Chemistry & Biology* 12.9 (2005), p. 1015–1028.

[78] Pierre-Jean L’HEUREUX, Julie CARREAU, Yoshua BENGIO, Olivier DELALLEAU et Shi Yi YUE. “Locally Linear Embedding for dimensionality reduction in QSAR”. In : *Journal of Computer-Aided Molecular Design* 18 (2004), p. 475–482.

[79] Yoshua BENGIO, Olivier DELALLEAU, Nicolas LE ROUX, Jean-Francois PAIEMENT, Pascal VINCENT et Marie OUIMET. “Learning eigenfunctions links spectral embedding and kernel PCA”. In : *Neural Computation* 16.10 (2004), p. 2197–2219.

[80] Yoshua BENGIO et Yves GRANDVALET. “No Unbiased Estimator of the Variance of K-Fold Cross-Validation”. In : *Journal of Machine Learning Research* 5 (2004), p. 1089–1105.

[81] R. COLLOBERT, Y. BENGIO et S. BENGIO. “Scaling Large Learning Problems with Hard Parallel Mixtures”. In : *International Journal of Pattern Recognition and Artificial Intelligence* 17.3 (2003), p. 349–365.

[82] Claude NADEAU et Yoshua BENGIO. “Inference for the Generalization Error”. In : *Machine Learning* 52.3 (2003), p. 239–281.

[83] J. GHOSN et Y. BENGIO. “Bias Learning, Knowledge Sharing”. In : *IEEE Transactions on Neural Networks* 14 (4 juil. 2003), p. 748–765.

[84] Yoshua BENGIO, Réjean DUCHARME, Pascal VINCENT et Christian JAUVIN. “A Neural Probabilistic Language Model”. In : *Journal of Machine Learning Research* 3 (2003), p. 1137–1155.

[85] Yoshua BENGIO et Nicolas CHAPADOS. “Extensions to Metric-Based Model Selection”. In : *Journal of Machine Learning Research* 3 (mar. 2003), p. 1209–1227.

[86] Ichiro TAKEUCHI, Yoshua BENGIO et Takafumi KANAMORI. “Robust Regression with Asymmetric Heavy-Tail Noise Distributions”. In : *Neural Computation* 14.10 (2002), p. 2469–2496.

[87] R. COLLOBERT, S. BENGIO et Y. BENGIO. “Parallel Mixture of SVMs for Very Large Scale Problems”. In : *Neural Computation* 14.5 (2002), p. 1105–1114.

[88] O. CHAPELLE, V. VAPNIK et Y. BENGIO. “Model Selection for Small-Sample Regression”. In : *Machine Learning Journal* 48.1 (2002), p. 9–23.

[89] P. VINCENT et Y. BENGIO. “Kernel Matching Pursuit”. In : *Machine Learning* 48 (2002), p. 165–187.

[90] Nicolas CHAPADOS et Yoshua BENGIO. “Cost Functions and Model Combination for VaR–based Asset Allocation using Neural Networks”. In : *IEEE Transactions on Neural Networks* 12.4 (2001), p. 890–906.

[91] Yoshua BENGIO, Vincent-Philippe LAUZON et Réjean DUCHARME. “Experiments on the Application of IOHMMs to Model Financial Returns Series”. In : *IEEE Transactions on Neural Networks* 12.1 (2001), p. 113–123.

[92] Holger SCHWENK et Yoshua BENGIO. “Boosting Neural Networks”. In : *Neural Computation* 12.8 (2000), p. 1869–1887.

[93] Yoshua BENGIO. “Gradient-Based Optimization of Hyperparameters”. In : *Neural Computation* 12.8 (2000), p. 1889–1900.

[94] S. BENGIO et Y. BENGIO. “Taking on the Curse of Dimensionality in Joint Distributions Using Neural Networks”. In : *IEEE Transactions on Neural Networks, special issue on Data Mining and Knowledge Discovery* 11.3 (2000), p. 550–557.

[95] Yoshua BENGIO. “Markovian Models for Sequential Data”. In : *Neural Computing Surveys* 2 (1999), p. 129–162.

[96] S. BENGIO, Y. BENGIO, J. ROBERT et G. BÉLANGER. “Stochastic Learning of Strategic Equilibria for Auctions”. In : *Neural Computation* 11.5 (1999), p. 1199–1209.

[97] L. BOTTOU, P. HAFFNER, P.G. HOWARD, P. SIMARD et Y. BENGIO. “High quality document image compression with DjVu”. In : *Journal of Electronic Imaging* 7.3 (1998), p. 410–425.

[98] Yann LECUN, Leon BOTTOU, Yoshua BENGIO et Patrick HAFFNER. “Gradient-Based Learning Applied to Document Recognition”. In : *Proceedings of the IEEE* 86.11 (nov. 1998), p. 2278–2324.

[99] Y. BENGIO, F. GINGRAS, B. GOULARD et J.-M. LINA. “Gaussian Mixture Densities for Classification of Nuclear Power Plant Data”. In : *Computers and Artificial Intelligence, special issue on Intelligent Technologies for Electric and Nuclear Power Plants* 17.2–3 (1998), p. 189–209.

[100] F. GINGRAS et Y. BENGIO. “Handling Asynchronous or Missing Financial Data with Recurrent Networks”. In : *International Journal of Computational Intelligence and Organizations* 1.3 (1998), p. 154–163.

[101] Yoshua BENGIO. “Using a Financial Training Criterion Rather than a Prediction Criterion”. In : *International Journal of Neural Systems* 8.4 (1997), p. 433–443.

[102] Y. BENGIO et P. FRASCONI. “Input/Output HMMs for Sequence Processing”. In : *IEEE Transactions on Neural Networks* 7.5 (1996), p. 1231–1249.

[103] Yoshua BENGIO et Paolo FRASCONI. “Diffusion of Context and Credit Information in Markovian Models”. In : *Journal of Artificial Intelligence Research* 3 (1995), p. 249–270.

[104] Yoshua BENGIO, Yann LECUN, Craig NOHL et Chris BURGES. “LeRec: A NN/HMM Hybrid for On-Line Handwriting Recognition”. In : *Neural Computation* 7.6 (1995), p. 1289–1303.

[105] S. BENGIO, Y. BENGIO et J. CLOUTIER. “On the search for new learning rules for ANNs”. In : *Neural Processing Letters* 2.4 (1995), p. 26–30.

[106] Y. BENGIO, P. SIMARD et P. FRASCONI. “Learning Long-Term Dependencies with Gradient Descent is Difficult”. In : *IEEE Transactions on Neural Networks* 5.2 (1994), p. 157–166.

[107] Yoshua BENGIO. “A Connectionist Approach to Speech Recognition”. In : *International Journal on Pattern Recognition and Artificial Intelligence* 7.4 (1993), p. 647–668.

[108] Yoshua BENGIO, Marco GORI et Renato DE MORI. “Learning the Dynamic Nature of Speech with Back-propagation for Sequences”. In : *Pattern Recognition Letters* 13.5 (1992), p. 375–385.

[109] Yoshua BENGIO, Renato DE MORI, Giovanni FLAMMIA et Ralf KOMPE. “Phonetically motivated acoustic parameters for continuous speech recognition using artificial neural networks”. In : *Speech Communication* 11.2–3 (1992), p. 261–271.

[110] Yoshua BENGIO, Renato DE MORI, Giovanni FLAMMIA et Ralf KOMPE. “Global Optimization of a Neural Network-Hidden Markov Model Hybrid”. In : *IEEE Transactions on Neural Networks* 3.2 (1992), p. 252–259.

[111] Yoshua BENGIO et Yannick POULIOT. “Efficient recognition of immunoglobulin domains from amino-acid sequences using a neural network”. In : *Computer Applications in the Biosciences* 6.2 (1990), p. 319–324.

[112] Piero COSI, Yoshua BENGIO et Renato DE MORI. “Phonetically-based multi-layered networks for acoustic property extraction and automatic speech recognition”. In : *Speech Communication* 9.1 (1990), p. 15–30.

[113] Y. BENGIO et R. De MORI. “Use of multilayer networks for the recognition of phonetic features and phonemes”. In : *Computational Intelligence* 5 (1989), p. 134–141.

[114] Yoshua BENGIO, Regis CARDIN, Renato DE MORI et Ettore MERLO. “Programmable execution of multi-layered networks for automatic speech recognition”. In : *Communications of the Association for Computing Machinery* 32.2 (1989), p. 195–199.

Articles in Refereed Conference Proceedings

[115] Minkai XU, Wujie WANG, Shitong LUO, Chence SHI, Yoshua BENGIO, Rafael GOMEZ-BOMBARELLI et Jian TANG. “An End-to-End Framework for Molecular Conformation Generation via Bilevel Programming”. In : *ICML 2021: 38th International Conference on Machine Learning*. 2021.

[116] Yoshua BENGIO, Prateek GUPTA, Tegan MAHARAJ, Nasim RAHAMAN, Martin WEISS, Tristan DELEU, Eilif MULLER, Meng QU, Victor SCHMIDT, Pierre-Luc ST-CHARLES et al. “Predicting infectiousness for proactive contact tracing”. In : *International Conference on Learning Representations*. 2021.

[117] Faruk AHMED, Yoshua BENGIO, Harm van SEIJEN et Aaron COURVILLE. “Systematic generalisation with group invariant predictions”. In : *International Conference on Learning Representations*. 2021.

[118] Joseph D VIVIANO, Becks SIMPSON, Francis DUTIL, Yoshua BENGIO et Joseph Paul COHEN. “Saliency is a possible red herring when diagnosing poor generalization”. In : *International Conference on Learning Representations*. 2021.

[119] François ST-HILAIRE et al. “A Comparative Study of Learning Outcomes for Online Learning Platforms”. In : *The 2021 conference on Artificial Intelligence in Education*. 2021, p. 331–337.

[120] Alex LAMB, Anirudh GOYAL, Agnieszka SŁOWIK, Michael MOZER, Philippe BEAUDOIN et Yoshua BENGIO. "Neural Function Modules with Sparse Arguments: A Dynamic Approach to Integrating Information across Layers". In : *International Conference on Artificial Intelligence and Statistics*. PMLR. 2021, p. 919–927.

[121] Rémi LE PRIOL, Reza BABANEZHAD, Yoshua BENGIO et Simon LACOSTE-JULIEN. "An analysis of the adaptation speed of causal models". In : *International Conference on Artificial Intelligence and Statistics*. PMLR. 2021, p. 775–783.

[122] B. SCHÖLKOPF, F. LOCATELLO, S. BAUER, N. R. KE, N. KALCHBRENNER, A. GOYAL et Y. BENGIO. "Toward Causal Representation Learning". In : *Proceedings of the IEEE* (2021), p. 1–23.

[123] Konrad ŻOŁNA, Chitwan SAHARIA, Leonard BOUSSIOUX, David Yu-Tung HUI, Maxime CHEVALIER-BOISVERT, Dzmitry BAHDANAU et Yoshua BENGIO. "Combating False Negatives in Adversarial Imitation Learning (Student Abstract)". In : *Proceedings of the AAAI Conference on Artificial Intelligence (Student Abstract)*. T. 34. 10. 2020, p. 13999–14000.

[124] Giancarlo KERG, Bhargav KANUPARTHI, Anirudh GOYAL, Kyle GOYETTE, Yoshua BENGIO et Guillaume LAJOIE. "Untangling tradeoffs between recurrence and self-attention in neural networks". In : *Advances in Neural Information Processing Systems, NeurIPS'2020*. T. 33. 2020, p. 19443–19454.

[125] Tong CHE, Ruixiang ZHANG, Jascha SOHL-DICKSTEIN, Hugo LAROCHELLE, Liam PAULL, Yuan CAO et Yoshua BENGIO. "Your gan is secretly an energy-based model and you should use discriminator driven latent sampling". In : *Advances in Neural Information Processing Systems, NeurIPS'2020*. T. 33. 2020, p. 12275–12287.

[126] Prateek GUPTA, Maxime GASSE, Elias B. KHALIL, Pawan Kumar MUDIGONDA, Andrea LODI et Yoshua BENGIO. "Hybrid Models for Learning to Branch". In : *Advances in Neural Information Processing Systems, NeurIPS'2020*. T. 33. 2020, p. 18087–18097.

[127] Yoshua BENGIO, Emma FREJINGER, Andrea LODI, Rahul PATEL et Sriram SANKARANARAYANAN. "A learning-based algorithm to quickly compute good primal solutions for Stochastic Integer Programs". In : *International Conference on Integration of Constraint Programming, Artificial Intelligence, and Operations Research*. Springer. 2020, p. 99–111.

[128] Valentin THOMAS, Fabian PEDREGOSA, Bart MERRIËNBOER, Pierre-Antoine MANZAGOL, Yoshua BENGIO et Nicolas LE ROUX. "On the interplay between noise and curvature and its effect on optimization and generalization". In : *International Conference on Artificial Intelligence and Statistics*. PMLR. 2020, p. 3503–3513.

[129] Ghouthi Boukli HACENE, Vincent GRIPON, Matthieu ARZEL, Nicolas FARRUGIA et Yoshua BENGIO. "Quantized Guided Pruning for Efficient Hardware Implementations of Deep Neural Networks". In : *2020 18th IEEE International New Circuits and Systems Conference (NEWCAS)*. 2020, p. 206–209.

[130] Nasim RAHAMAN, Steffen WOLF, Anirudh GOYAL, Roman REMME et Yoshua BENGIO. "Learning the Arrow of Time for Problems in Reinforcement Learning". In : *ICLR'2020* (2020).

[131] Yoshua BENGIO, Tristan DELEU, Nasim RAHAMAN, Rosemary KE, Sébastien LACHAPELLE, Olexa BILANIUK, Anirudh GOYAL et Christopher PAL. "A meta-transfer objective for learning to disentangle causal mechanisms". In : *ICLR'2020*. 2020.

[132] Anirudh GOYAL, Shagun SODHANI, Jonathan BINAS, Xue Bin PENG, Sergey LEVINE et Yoshua BENGIO. "Reinforcement Learning with Competitive Ensembles of Information-Constrained Primitives". In : *ICLR'2020*. 2020.

[133] Anirudh GOYAL, Yoshua BENGIO, Matthew BOTVINICK et Sergey LEVINE. "The variational bandwidth bottleneck: Stochastic evaluation on an information budget". In : *ICLR'2020*. 2020.

[134] Boris N ORESHKIN, Dmitri CARPOV, Nicolas CHAPADOS et Yoshua BENGIO. "N-BEATS: Neural basis expansion analysis for interpretable time series forecasting". In : *ICLR'2020*. 2020.

[135] Mirco RAVANELLI, Jianyuan ZHONG, Santiago PASCUAL, Pawel SWIETOJANSKI, Joao MONTEIRO, Jan TRMAL et Yoshua BENGIO. "Multi-task self-supervised learning for Robust Speech Recognition". In : *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2020, p. 6989–6993.

[136] Surya Karthik MUKKAVILLI, Yimeng MIN, Ata MADANCHI, Vitoria Barin PACELA, Shivam PATEL et Yoshua BENGIO. "Integrated Climate Extremes: Modeling Future Impacts for Visualizing Climate Change". In : *100th American Meteorological Society Annual Meeting*. AMS. 2020.

[137] Y MIN, SK MUKKAVILLI et Y BENGIO. "IceNet: A large-scale dataset for tracking ice flow using unsupervised learning with adversarial networks". In : *100th American Meteorological Society Annual Meeting*. AMS. 2020.

[138] Sherjil OZAIR, Corey LYNCH, Yoshua BENGIO, Aaron VAN DEN OORD, Sergey LEVINE et Pierre SERMANET. "Wasserstein dependency measure for representation learning". In : *Advances in Neural Information Processing Systems, NeurIPS'2019*. 2019, p. 15578–15588.

[139] Taesup KIM, Sungjin AHN et Yoshua BENGIO. "Variational Temporal Abstraction". In : *Advances in Neural Information Processing Systems, NeurIPS'2019*. 2019, p. 11566–11575.

[140] Giancarlo KERG, Kyle GOYETTE, Maximilian Puelma TOUZEL, Gauthier GIDEL, Eugene VORONTSOV, Yoshua BENGIO et Guillaume LAJOIE. "Non-normal Recurrent Neural Network (nnRNN): learning long time dependencies while improving expressivity with transient dynamics". In : *Advances in Neural Information Processing Systems, NeurIPS'2019*. 2019, p. 13591–13601.

[141] Rahaf ALJUNDI, Min LIN, Baptiste GOUJAUD et Yoshua BENGIO. "Gradient based sample selection for online continual learning". In : *Advances in Neural Information Processing Systems, NeurIPS'2019*. 2019, p. 11816–11825.

[142] Ankesh ANAND, Evan RACAH, Sherjil OZAIR, Yoshua BENGIO, Marc-Alexandre CÔTÉ et R Devon HJELM. "Unsupervised state representation learning in atari". In : *Advances in Neural Information Processing Systems, NeurIPS'2019*. 2019, p. 8766–8779.

[143] Maxence ERNOULT, Julie GROLLIER, Damien QUERLIOZ, Yoshua BENGIO et Benjamin SCELLIER. "Updates of Equilibrium Prop Match Gradients of Backprop Through Time in an RNN with Static Input". In : *Advances in Neural Information Processing Systems, NeurIPS'2019*. 2019, p. 7079–7089.

[144] Kundan KUMAR, Rithesh KUMAR, Thibault de BOISSIERE, Lucas GESTIN, Wei Zhen TEOH, Jose SOTELO, Alexandre de BRÉBISSON, Yoshua BENGIO et Aaron C COURVILLE. "Melgan: Generative adversarial networks for conditional waveform synthesis". In : *Advances in Neural Information Processing Systems, NeurIPS'2019*. 2019, p. 14881–14892.

[145] Christopher BECKHAM, Sina HONARI, Vikas VERMA, Alex M LAMB, Farnoosh GHADIRI, R Devon HJELM, Yoshua BENGIO et Chris PAL. "On Adversarial Mixup Resynthesis". In : *Advances in Neural Information Processing Systems, NeurIPS'2019*. 2019, p. 4348–4359.

[146] Devansh ARPIT, Viéctor CAMPOS et Yoshua BENGIO. “How to initialize your network? robust initialization for weightnorm & resnets”. In : *Advances in Neural Information Processing Systems, NeurIPS’2019*. 2019, p. 10900–10909.

[147] Alaaeldin EL-NOUBY, Shikhar SHARMA, Hannes SCHULZ, Devon HJELM, Layla El ASRI, Samira Ebrahimi KAHOU, Yoshua BENGIO et Graham W TAYLOR. “Tell, draw, and repeat: Generating and modifying images based on continual linguistic instruction”. In : *Proceedings of the IEEE International Conference on Computer Vision, ICCV’2019*. 2019, p. 10304–10312.

[148] Vincent FRANÇOIS-LAVET, Yoshua BENGIO, Doina PRECUP et Joelle PINEAU. “Combined reinforcement learning via abstract representations”. In : *AAAI’2019*. T. 33. 2019, p. 3582–3589.

[149] Kyle KASTNER, Jo∼ao Felipe SANTOS, Yoshua BENGIO et Aaron COURVILLE. “Representation Mixing for TTS Synthesis”. In : *ICASSP’2019*. 2019, p. 5906–5910.

[150] Sanghyun YOO, Inchul SONG et Yoshua BENGIO. “A Highly Adaptive Acoustic Model for Accurate Multi-dialect Speech Recognition”. In : *ICASSP’2019*. 2019, p. 5716–5720.

[151] Jessica AF THOMPSON, Marc SCHÖNWIESNER, Yoshua BENGIO et Daniel WILLETT. “How transferable are features in convolutional neural network acoustic models across languages?” In : *ICASSP’2019*. 2019, p. 2827–2831.

[152] Saeid Asgari TAGHANAKI, Mohammad HAVAEI, Tess BERTHIER, Francis DUTIL, Lisa DI JORIO, Ghassan HAMARNEH et Yoshua BENGIO. “InfoMask: Masked Variational Latent Representation to Localize Chest Disease”. In : *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2019, p. 739–747.

[153] Sarath CHANDAR, Chinnadhurai SANKAR, Eugene VORONTSOV, Samira Ebrahimi KAHOU et Yoshua BENGIO. “Towards non-saturating recurrent units for modelling long-term dependencies”. In : *Proceedings of the AAAI Conference on Artificial Intelligence*. T. 33. 2019, p. 3280–3287.

[154] Alex LAMB, Vikas VERMA, Juho KANNALA et Yoshua BENGIO. “Interpolated Adversarial Training: Achieving Robust Neural Networks Without Sacrificing Too Much Accuracy”. In : *Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security*. 2019, p. 95–103.

[155] Md Mahfuzur Rahman SIDDIQUEE, Zongwei ZHOU, Nima TAJBAKHSH, Ruibin FENG, Michael B GOTWAY, Yoshua BENGIO et Jianming LIANG. “Learning fixed points in generative adversarial networks: From image-to-image translation to disease detection and localization”. In : *Proceedings of the IEEE International Conference on Computer Vision, ICCV’2019*. 2019, p. 191–200.

[156] Alex LAMB, Jonathan BINAS, Anirudh GOYAL, Sandeep SUBRAMANIAN, Ioannis MITLIAGKAS, Yoshua BENGIO et Michael MOZER. “State-Reification Networks: Improving Generalization by Modeling the Distribution of Hidden Representations”. In : *ICML’2019*. 2019, p. 3622–3631.

[157] Nasim RAHAMAN, Aristide BARATIN, Devansh ARPIT, Felix DRAXLER, Min LIN, Fred HAMPRECHT, Yoshua BENGIO et Aaron COURVILLE. “On the Spectral Bias of Neural Networks”. In : *ICML’2019*. 2019, p. 5301–5310.

[158] Meng QU, Yoshua BENGIO et Jian TANG. “GMNN: Graph Markov Neural Networks”. In : *ICML’2019*. 2019, p. 5241–5250.

[159] Vikas VERMA, Alex LAMB, Christopher BECKHAM, Amir NAJAFI, Ioannis MITLIAGKAS, David LOPEZ-PAZ et Yoshua BENGIO. “Manifold Mixup: Better Representations by Interpolating Hidden States”. In : *ICML’2019*. 2019, p. 6438–6447.

[160] Maxime CHEVALIER-BOISVERT, Dzmitry BAHDANAU, Salem LAHLOU, Lucas WILLEMS, Chitwan SAHARIA, Thien Huu NGUYEN et Yoshua BENGIO. "BabyAI: First steps towards grounded language learning with a human in the loop". In : *ICLR'2019*. 2019.

[161] Nan Rosemary KE, Amanpreet SINGH, Ahmed TOUATI, Anirudh GOYAL, Yoshua BENGIO, Devi PARIKH et Dhruv BATRA. "Modeling the Long Term Future in Model-Based Reinforcement Learning". In : *ICLR'2019*. 2019.

[162] Anirudh GOYAL, Philemon BRAKEL, William FEDUS, Soumye SINGHAL, Timothy LILLICRAP, Sergey LEVINE, Hugo LAROCHELLE et Yoshua BENGIO. "Recall Traces: Backtracking Models for Efficient Reinforcement Learning". In : *ICLR'2019*. 2019.

[163] Anirudh GOYAL, Riashat ISLAM, Daniel STROUSE, Zafarali AHMED, Matthew BOTVINICK, Hugo LAROCHELLE, Sergey LEVINE et Yoshua BENGIO. "Infobot: Transfer and exploration via the information bottleneck". In : *ICLR'2019*. 2019.

[164] R Devon HJELM, Alex FEDOROV, Samuel LAVOIE-MARCHILDON, Karan GREWAL, Phil BACHMAN, Adam TRISCHLER et Yoshua BENGIO. "Learning deep representations by mutual information estimation and maximization". In : *ICLR'2019*. 2019.

[165] Ali FARSHCHIAN, Juan A GALLEGO, Joseph P COHEN, Yoshua BENGIO, Lee E MILLER et Sara A SOLLA. "Adversarial Domain Adaptation for Stable Brain-Machine Interfaces". In : *ICLR'2019*. 2019.

[166] Petar VELIČKOVIĆ, William FEDUS, William L HAMILTON, Pietro LIÒ, Yoshua BENGIO et R Devon HJELM. "Deep Graph Infomax". In : *ICLR'2019*. 2019.

[167] Alexandre PICHÉ, Valentin THOMAS, Cyril IBRAHIM, Yoshua BENGIO et Chris PAL. "Probabilistic planning with sequential monte carlo methods". In : *ICLR'2019*. 2019.

[168] Mariya TONEVA, Alessandro SORDONI, Remi COMBES, Adam TRISCHLER, Yoshua BENGIO et Geoffrey GORDON. "An Empirical Study of Example Forgetting during Deep Neural Network Learning". In : *ICLR'2019*. 2019.

[169] Bhargav KANUPARTHI, Devansh ARPIT, Giancarlo KERG, Nan Rosemary KE, Ioannis MITLIAGKAS et Yoshua BENGIO. "h-detach: Modifying the LSTM Gradient Towards Better Optimization". In : *ICLR'2019*. 2019.

[170] Stanisław JASTRZĘBSKI, Zachary KENTON, Nicolas BALLAS, Asja FISCHER, Yoshua BENGIO et Amos STORKEY. "On the relation between the sharpest directions of DNN loss and the SGD step length". In : *ICLR'2019*. 2019.

[171] Mirco RAVANELLI, Titouan PARCOLLET et Yoshua BENGIO. "The PyTorch-Kaldi Speech Recognition Toolkit". In : *ICASSP'2019*. 2019.

[172] Homanga BHARADHWAJ, Zihan WANG, Yoshua BENGIO et Liam PAULL. "A data-efficient framework for training and sim-to-real transfer of navigation policies". In : *2019 International Conference on Robotics and Automation (ICRA)*. IEEE. 2019, p. 782–788.

[173] A. GONZALEZ-GARCIA, J. VAN DE WEIJER et Y. BENGIO. "Image-to-image translation for cross-domain disentanglement". In : *NeurIPS'2018*.

[174] Ruixiang ZHANG, Tong CHE, Zoubin GHAHRAMANI, Yoshua BENGIO et Yangqiu SONG. "MetaGAN: An Adversarial Approach to Few-Shot Learning". In : *NeurIPS'2018*. 2018.

[175] T. KIM, J. YOON, O. DIA, S. KIM, Y. BENGIO et S. AHN. "Bayesian Model-Agnostic Meta-Learning". In : *NeurIPS'2018*. 2018.

[176] N. R. KE, A. GOYAL, O. BILANIUK, J. BINAS, M. C. MOZER, C. PAL et Y. BENGIO. "Sparse Attentive Backtracking: Temporal CreditAssignment Through Reminding". In : *NeurIPS'2018*. 2018.

[177] J. SACRAMENTO, R. PONTE COSTA, Y. BENGIO et W. SENN. "Dendritic error backpropagation in deep cortical microcircuits". In : *NeurIPS'2018*. 2018.

[178] Mohamed BELGHAZI, Aristide BARATIN, Sai RAJESWAR, Sherjil OZAIR, Yoshua BENGIO, Devon HJELM et Aaron COURVILLE. "Mutual Information Neural Estimation". In : *ICML'2018*. 2018.

[179] Nan KE, Konrad ZOLNA, Alessandro SORDONI, Zhouhan LIN, Adam TRISCHLER, Yoshua BENGIO, Joelle PINEAU, Laurent CHARLIN et Christopher PAL. "Focused Hierarchical RNNs for Conditional Sequence Processing". In : *ICML'2018*. 2018.

[180] Florian BORDES, Tess BERTHIER, Lisa Di JORIO, Pascal VINCENT et Yoshua BENGIO. "Iteratively unveiling new regions of interest in Deep Learning models". In : *Medical Imaging with Deep Learning, MIDL'2018*. 2018.

[181] Mirco RAVANELLI, Dmitriy SERDYUK et Yoshua BENGIO. "Twin regularization for online speech recognition". In : *Interspeech*. 2018.

[182] Titouan PARCOLLET, Ying ZHANG, Chiheb TRABELSI, Mohamed MORCHID, Renato de MORI, Georges LINARES et Yoshua BENGIO. "Quaternion Convolutional Neural Networks for End-to-End Automatic Speech Recognition". In : *Interspeech*. 2018.

[183] Stanislaw JASTRZĘBSKI, Zachary KENTON, Devansh ARPIT, Nicolas BALLAS, Asja FISCHER, Yoshua BENGIO et Amos STORKEY. "Width of Minima Reached by Stochastic Gradient Descent is Influenced by Learning Rate to Batch Size Ratio". In : *International Conference on Artificial Neural Networks*. Springer. 2018, p. 392–402.

[184] R Devon HJELM, Athul P JACOB, Adam TRISCHLER, Tong CHE, Kyunghyun CHO et Yoshua BENGIO. "Boundary Seeking GANs". In : *ICLR'2018*. 2018.

[185] Sandeep SUBRAMANIAN, Adam TRISCHLER, Yoshua BENGIO et Christopher PAL. "Learning General Purpose Distributed Sentence Representations via Large Scale Multi-task Learning". In : *ICLR'2018*. 2018.

[186] Dmitriy SERDYUK, Nan Rosemary KE, Alessandro SORDONI, Adam TRISCHLER, Christopher PAL et Yoshua BENGIO. "Twin Networks: Matching the Future for Sequence Generation". In : *ICLR'2018*. 2018.

[187] Konrad ZOLNA, Devansh ARPIT, Dendi SUHUBDY et Yoshua BENGIO. "Fraternal Dropout". In : *ICLR'2018*. 2018.

[188] Chiheb TRABELSI, Olexa BILANIUK, Ying ZHANG, Dmitriy SERDYUK, Sandeep SUBRAMANIAN, Joao Felipe SANTOS, Soroush MEHRI, Negar ROSTAMZADEH, Yoshua BENGIO et Christopher PAL. "Deep Complex Networks". In : *ICLR'2018*. 2018.

[189] Petar VELIČKOVIĆ, Guillem Cucurull PREIXENS, Arantxa Casanova PAGA, Adriana ROMERO, Pietro LIÒ et Yoshua BENGIO. "Graph Attention Networks". In : *ICLR'2018*. 2018.

[190] Stanislaw JASTRZEBSKI, Devansh ARPIT, Nicolas BALLAS, Vikas VERMA, Tong CHE et Yoshua BENGIO. "Residual Connections Encourage Iterative Inference". In : *ICLR'2018*. 2018.

[191] Inchul SONG, Junyoung CHUNG, Taesup KIM et Yoshua BENGIO. "Dynamic Frame Skipping for Fast Speech Recognition in Recurrent Neural Network Based Acoustic Models". In : *ICASSP'2018*. 2018, p. 4984–4988.

[192] SI MIMILAKIS, K DROSSOS, JF SANTOS, G SCHULLER, T VIRTANEN et Y BENGIO. "Monaural singing voice separation with skip-filtering connections and recurrent inference of time-frequency mask". In : *Proc. ICASSP'2018*. 2018.

[193] Dmitriy SERDYUK, Yongqiang WANG, Christian FUEGEN, Anuj KUMAR, Baiyang LIU et Yoshua BENGIO. "Towards end-to-end spoken language understanding". In : *ICASSP'2018*. 2018, p. 5754–5758.

[194] Konstantinos DROSSOS, Stylianos Ioannis MIMILAKIS, Dmitriy SERDYUK, Gerald SCHULLER, Tuomas VIRTANEN et Yoshua BENGIO. "MaD TwinNet: Masker-denoiser architecture with twin networks for monaural sound source separation". In : *IJCNN'2018*. 2018, p. 1–8.

[195] Çaglar GÜLÇEHRE, Francis DUTIL, Adam TRISCHLER et Yoshua BENGIO. "Plan, Attend, Generate: Planning for Sequence-to-Sequence Models". In : *NIPS'2017*. 2017, p. 5480–5489.

[196] Alex LAMB, Devon R HJELM, Yaroslav GANIN, Joseph Paul COHEN, Aaron COURVILLE et Yoshua BENGIO. "GibbsNet: Iterative Adversarial Inference for Deep Graphical Models". In : *NIPS'2017*. 2017, p. 5095–5104.

[197] Anirudh GOYAL, Alessandro SORDONI, Marc-Alexandre CÔTÉ, Nan Rosemary KE et Yoshua BENGIO. "Z-Forcing: Training Stochastic Recurrent Networks". In : *NIPS'2017*. 2017, p. 6716–6726.

[198] Iulian V. SERBAN, Alessandro SORDONI, Ryan LOWE, Laurent CHARLIN, Joelle PINEAU, Aaron COURVILLE et Yoshua BENGIO. "A Hierarchical Latent Variable Encoder-Decoder Model for Generating Dialogues". In : *AAAI'2017*. 2017, p. 1583.

[199] Laurent DINH, Razvan PASCANU, Samy BENGIO et Yoshua BENGIO. "Sharp Minima Can Generalize For Deep Nets". In : *ICML'2017*. T. 70. Août 2017, p. 1019–1028.

[200] David KRUEGER, Nicolas BALLAS, Stanislaw JASTRZEBSKI, Devansh ARPIT, Maxinder S. KANWAL, Tegan MAHARAJ, Emmanuel BENGIO, Asja FISCHER, Aaron COURVILLE, Simon LACOSTE-JULIEN et Yoshua BENGIO. "A Closer Look at Memorization in Deep Networks". In : *ICML'2017*. 2017.

[201] Mirco RAVANELLI, Philemon BRAKEL, Maurizio OMOLOGO et Yoshua BENGIO. "A network of deep neural networks for distant speech recognition". In : *ICASSP'2017*. 1358. 2017.

[202] Taesup KIM, Inchul SONG et Yoshua BENGIO. "Dynamic Layer Normalization for Adaptive Neural Acoustic Modeling in Speech Recognition". In : *Interspeech 2017*. Août 2017.

[203] Ryan LOWE, Michael NOSEWORTHY, Iulian V. SERBAN, Nicolas ANGELARD-GONTIER, Yoshua BENGIO et Joelle PINEAU. "Towards an Automatic Turing Test: Learning to Evaluate Dialogue Responses". In : *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*. 2017.

[204] Anh NGUYEN, Jason YOSINSKI, Yoshua BENGIO, Alexey DOSOVITSKIY et Jeff CLUNE. "Plug & play generative networks: Conditional iterative generation of images in latent space". In : *CVPR'2017*. 2017.

[205] Caglar GÜLÇEHRE, Marcin MOCZULSKI, Francesco VISIN et Yoshua BENGIO. "Mollifying Networks". In : *ICLR'2017*. 2017.

[206] David WARDE-FARLEY et Yoshua BENGIO. "Improving Generative Adversarial Networks With Denoising Feature Matching". In : *ICLR'2017*. 2017.

[207] Iulian V. SERBAN, Tim KLINGER, Gerald TESAURO, Kartik TALAMADUPULA, Bowen ZHOU, Yoshua BENGIO et Aaron COURVILLE. "Multiresolution Recurrent Neural Networks: An Application to Dialogue Response Generation". In : *AAAI'2017*. 2017, p. 1641.

[208] Dzmitry BAHDANAU, Philemon BRAKEL, Kelvin XU, Anirudh GOYAL, Ryan LOWE, Joelle PINEAU, Aaron COURVILLE et Yoshua BENGIO. "An Actor-Critic Algorithm for Sequence Prediction". In : *ICLR'2017*. 2017.

[209] Junyoung CHUNG, Sungjin AHN et Yoshua BENGIO. "Hierarchical Multiscale Recurrent Neural Networks". In : *ICLR'2017*. 2017.

[210] David KRUEGER, Tegan MAHARAJ, János KRAMÁR, Mohammad PEZESHKI, Nicolas BALLAS, Nan Rosemary KE, Anirudh GOYAL, Yoshua BENGIO, Aaron COURVILLE et Christopher PAL. "Zoneout: Regularizing RNNs by Randomly Preserving Hidden Activations". In : *ICLR'2017*. 2017.

[211] Zhouhan LIN, Minwei FENG, Cicero Nogueira dos SANTOS, Mo YU, Bing XIANG, Bowen ZHOU et Yoshua BENGIO. "A Structured Self-Attentive Sentence Embedding". In : *ICLR'2017*. 2017.

[212] Soroush MEHRI, Kundan KUMAR, Ishaan GULRAJANI, Rithesh KUMAR, Shubham JAIN, Jose SOTELO, Aaron COURVILLE et Yoshua BENGIO. "SampleRNN: An Unconditional End-to-End Neural Audio Generation Model". In : *ICLR'2017*. 2017.

[213] Adriana ROMERO, Pierre-Luc CARRIER, Akram ERRAQABI, Tristan SYLVAIN, Alex AUVOLAT, Etienne DEJOIE, Marc-André LEGAULT, Marie-Pierre DUBÉ, Julie G. HUSSIN et Yoshua BENGIO. "Diet Networks: Thin Parameters for Fat Genomic". In : *ICLR'2017*. 2017.

[214] Alex LAMB, Anirudh GOYAL, Ying ZHANG, Saizheng ZHANG, Aaron COURVILLE et Yoshua BENGIO. "Professor Forcing: A New Algorithm for Training Recurrent Networks". In : *NIPS'2016*. 2016.

[215] Martin ARJOVSKY, Amar SHAH et Yoshua BENGIO. "Unitary Evolution Recurrent Neural Networks". In : *ICML'2016*. 2016.

[216] Yuhuai WU, Saizheng ZHANG, Ying ZHANG, Yoshua BENGIO et Ruslan SALAKHUTDINOV. "On Multiplicative Integration with Recurrent Neural Networks". In : *NIPS'2016*. 2016.

[217] Saizheng ZHANG, Yuhuai WU, Tong CHE, Zhouhan LIN, Roland MEMISEVIC, Ruslan SALAKHUTDINOV et Yoshua BENGIO. "Architectural Complexity Measures of Recurrent Neural Networks". In : *NIPS'2016*. 2016.

[218] Itay HUBARA, Matthieu COURBARIAUX, Daniel SOUDRY, Ran EL-YANIV et Yoshua BENGIO. "Binarized Neural Networks". In : *NIPS'2016*. 2016.

[219] Junyoung CHUNG, Kyunghyun CHO et Yoshua BENGIO. "NYU-MILA Neural Machine Translation Systems for WMT'16". In : *First Conference on Machine Translation*. 2016.

[220] Ying ZHANG, Mohammad PEZESHKI, Philemon BRAKEL, Saizheng ZHANG, César LAURENT, Yoshua BENGIO et Aaron COURVILLE. "Towards End-to-End Speech Recognition with Deep Convolutional Neural Networks". In : *Interspeech 2016*. 2016, p. 410–414.

[221] Mirco RAVANELLI, Philemon BRAKEL, Maurizio OMOLOGO et Yoshua BENGIO. "Batch-normalized joint training for DNN-based distant speech recognition". In : *2016 IEEE Spoken Language Technology Workshop (SLT)*. Déc. 2016, p. 28–34.

[222] Mohammad HAVAEI, Nicolas GUIZARD, Nicolas CHAPADOS et Yoshua BENGIO. "HeMIS: Hetero-Modal Image Segmentation". In : *Int. Conf. Medical Image Computing and Computer-Assisted Intervention, MICCAI-2016*. 2016, p. 469–477.

[223] Julian SERBAN, Alessandro SORDONI, Yoshua BENGIO, Aaron COURVILLE et Joelle PINEAU. "Building end-to-end dialogue systems using generative hierarchical neural network models". In : *AAAI'2016*. 2016.

[224] Ying ZHANG, Mohammad PEZESHKI, Philémon BRAKEL, Saizheng ZHANG, César LAURENT, Yoshua BENGIO et Aaron COURVILLE. "Towards End-to-End Speech Recognition with Deep Convolutional Neural Networks". In : *Interspeech'2016*. 2016, p. 410–414.

[225] Dzmitry Bahdanau Jan CHOROWSKI, Dmitriy SERDYUK, Philemon BRAKEL et Yoshua BENGIO. “End-to-end attention-based large vocabulary speech recognition”. In : *ICASSP’2016*. 2016, p. 4945–4949.

[226] César LAURENT, Gabriel PEREYRA, Philemon BRAKEL, Ying ZHANG et Yoshua BENGIO. “Batch normalized recurrent neural networks”. In : *ICASSP’2016*. 2016, p. 2657–2661.

[227] Jorg BORNSCHEIN, Samira SHABANIAN, Asja FISCHER et Yoshua BENGIO. “Training Bidirectional Helmholtz Machines”. In : *ICML’2016*. 2016.

[228] Caglar GÜLÇEHRE, Marcin MOCZULSKI, Misha DENIL et Yoshua BENGIO. “Noisy Activation Functions”. In : *ICML’2016*. 2016.

[229] Mohammad PEZESHKI, Linxi FAN, Philemon BRAKEL, Aaron COURVILLE et Yoshua BENGIO. “Deconstructing the Ladder Network Architecture”. In : *ICML’2016*. 2016.

[230] Zhouhan LIN, Matthieu COURBARIAUX, Roland MEMISEVIC et Yoshua BENGIO. “Neural Networks with Few Multiplications”. In : *ICLR’2016*. 2016.

[231] Junyoung CHUNG, Kyle KASTNER, Laurent DINH, Kratarth GOEL, Aaron COURVILLE et Yoshua BENGIO. “A Recurrent Latent Variable Model for Sequential Data”. In : *NIPS’2015*. 2015.

[232] Jan CHOROWSKI, Dzmitry BAHDANAU, Dmitriy SERDYUK, Kyunghyun CHO et Yoshua BENGIO. “Attention-Based Models for Speech Recognition”. In : *NIPS’2015*. 2015.

[233] Caglar GULCEHRE, Kyunghyun CHO, Razvan PASCANU et Yoshua BENGIO. “Learned-norm pooling for deep feedforward and recurrent neural networks”. In : *Machine Learning and Knowledge Discovery in Databases (ECML/PKDD)*. 2015.

[234] Matthieu COURBARIAUX, Yoshua BENGIO et Jean-Pierre DAVID. “BinaryConnect: Training Deep Neural Networks with binary weights during propagations”. In : *NIPS’2015*. 2015.

[235] Yann DAUPHIN, Harm de VRIES et Yoshua BENGIO. “Equilibrated adaptive learning rates for non-convex optimization”. In : *NIPS’2015*. 2015.

[236] Alessandro SORDONI, Yoshua BENGIO, Hossein VAHABI, Christina LIOMA, Jakob GRUE SIMONSEN et Jian-Yun NIE. “A hierarchical recurrent encoder-decoder for generative context-aware query suggestion”. In : *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*. 2015, p. 553–562.

[237] Sébastien JEAN, Orhan FIRAT, Kyunghyun CHO, Roland MEMISEVIC et Yoshua BENGIO. “Montreal Neural Machine Translation Systems for WMT15”. In : *Proceedings of the Tenth Workshop on Statistical Machine Translation*. 2015, p. 134–140.

[238] Sébastien JEAN, Kyunghyun CHO, Roland MEMISEVIC et Yoshua BENGIO. “On using very large target vocabulary for neural machine translation”. In : *ACL-IJCNLP’2015*. 2015.

[239] Junyoung CHUNG, Caglar GÜLÇEHRE, Kyunghyun CHO et Yoshua BENGIO. “Gated Feedback Recurrent Neural Networks”. In : *ICML’2015*. 2015, p. 2067–2075.

[240] Dong-Hyun LEE, Saizheng ZHANG, Asja FISCHER et Yoshua BENGIO. “Difference Target Propagation”. In : *Machine Learning and Knowledge Discovery in Databases (ECML/PKDD)*. 2015.

[241] Jörg BORNSCHEIN et Yoshua BENGIO. “Reweighted Wake-Sleep”. In : *ICLR’2015, arXiv:1406.2751*. 2015.

[242] Dzmitry BAHDANAU, Kyunghyun CHO et Yoshua BENGIO. “Neural Machine Translation by Jointly Learning to Align and Translate”. In : *ICLR’2015, arXiv:1409.0473*. 2015.

[243] Adriana ROMERO, Nicolas BALLAS, Samira EBRAHIMI KAHOU, Antoine CHASSANG, Carlo GATTA et Yoshua BENGIO. “FitNets: Hints for Thin Deep Nets”. In : *ICLR’2015, arXiv:1412.6550*. 2015.

[244] Grégoire MESNIL, Salah RIFAI, Antoine BORDES, Xavier GLOROT, Yoshua BENGIO et Pascal VINCENT. “Unsupervised Learning of Semantics of Object Detections for Scene Categorization”. In : *Pattern Recognition Applications and Methods, Advances in Intelligent Systems and Computing*. Sous la dir. d’A. FRED et M. DE MARSICO. Springer International Publishing Switzerland, 2015, p. 209–224.

[245] Li YAO, Sherjil OZAIR, Kyunghyun CHO et Yoshua BENGIO. “On the Equivalence between Deep NADE and Generative Stochastic Networks”. In : *Machine Learning and Knowledge Discovery in Databases*. 2014.

[246] Tapani RAIKO, Li YAO, Kyunghyun CHO et Yoshua BENGIO. “Iterative Neural Autoregressive Distribution Estimator (NADE-k)”. In : *NIPS’2014*. 2014.

[247] Jason YOSINSKI, Jeff CLUNE, Yoshua BENGIO et Hod LIPSON. “How transferable are features in deep neural networks?” In : *NIPS’2014*. 2014.

[248] Ian J. GOODFELLOW, Jean POUGET-ABADIE, Mehdi MIRZA, Bing XU, David WARDE-FARLEY, Sherjil OZAIR, Aaron COURVILLE et Yoshua BENGIO. “Generative Adversarial Networks”. In : *NIPS’2014*. 2014.

[249] Guido F. MONTUFAR, Razvan PASCANU, KyungHyun CHO et Yoshua BENGIO. “On the Number of Linear Regions of Deep Neural Networks”. In : *NIPS’2014*. 2014.

[250] Yann DAUPHIN, Razvan PASCANU, Caglar GULCEHRE, Kyunghyun CHO, Surya GANGULI et Yoshua BENGIO. “Identifying and attacking the saddle point problem in high-dimensional non-convex optimization”. In : *NIPS’2014*. 2014.

[251] Kyunghyun CHO, Bart van MERRIENBOER, Caglar GULCEHRE, Fethi BOUGARES, Holger SCHWENK et Yoshua BENGIO. “Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation”. In : *Proceedings of the Empiricial Methods in Natural Language Processing (EMNLP 2014)*. Oct. 2014.

[252] Alessandro SORDONI, Yoshua BENGIO et Jian-Yun NIE. “Learning Concept Embeddings for Query Expansion by Quantum Entropy Minimization”. In : *AAAI*. 2014, p. 1586–1592.

[253] Yoshua BENGIO, Eric THIBODEAU-LAUFER et Jason YOSINSKI. “Deep Generative Stochastic Networks Trainable by Backprop”. In : *ICML’2014*. 2014.

[254] Minmin CHEN, Kilian WEINBERGER, Fei SHA et Yoshua BENGIO. “Marginalized Denoising Auto-encoders for Nonlinear Representations”. In : *ICML’2014*. 2014.

[255] David WARDE-FARLEY, Ian J. GOODFELLOW, Aaron COURVILLE et Yoshua BENGIO. “An empirical analysis of dropout in piecewise linear networks”. In : *ICLR’2014*. 2014.

[256] Ian J. GOODFELLOW, Mehdi MIRZA, Da XIAO, Aaron COURVILLE et Yoshua BENGIO. “An Empirical Investigation of Catastrophic Forgeting in Gradient-Based Neural Networks”. In : *ICLR’2014*. 2014.

[257] Yoshua BENGIO, Li YAO et Kyunghyun CHO. “Bounding the Test Log-Likelihood of Generative Models”. In : *ICLR’2014*. 2014.

[258] Razvan PASCANU, Caglar GÜLÇEHRE, Kyunghyun CHO et Yoshua BENGIO. “How to Construct Deep Recurrent Neural Networks”. In : *ICLR’2014*. 2014.

[259] Razvan PASCANU, Guido MONTUFAR et Yoshua BENGIO. “On the number of inference regions of deep feed forward networks with piece-wise linear activations”. In : *ICLR’2014*. 2014.

[260] Razvan PASCANU et Yoshua BENGIO. “Revisiting Natural Gradient for Deep Networks”. In : *ICLR’2014*. 2014.

[261] Ian J. GOODFELLOW, Mehdi MIRZA, Aaron COURVILLE et Yoshua BENGIO. “Multi-Prediction Deep Boltzmann Machines”. In : *Advances in Neural Information Processing Systems 26 (NIPS 2013)*. NIPS Foundation (http://books.nips.cc), déc. 2013.

[262] Yann DAUPHIN et Yoshua BENGIO. “Stochastic Ratio Matching of RBMs for Sparse High-Dimensional Inputs”. In : *Advances in Neural Information Processing Systems 26 (NIPS 2013)*. NIPS Foundation (http://books.nips.cc), 2013.

[263] Yoshua BENGIO, Li YAO, Guillaume ALAIN et Pascal VINCENT. “Generalized Denoising Auto-Encoders as Generative Models”. In : *NIPS’2013*. 2013.

[264] Nicolas BOULANGER-LEWANDOWSKI, Yoshua BENGIO et Pascal VINCENT. “High-dimensional Sequence Transduction”. In : *Proc. ICASSP 3*. 2013.

[265] Ian J. GOODFELLOW, David WARDE-FARLEY, Mehdi MIRZA, Aaron COURVILLE et Yoshua BENGIO. “Maxout Networks”. In : *ICML’2013*. 2013.

[266] Razvan PASCANU, Tomas MIKOLOV et Yoshua BENGIO. “On the difficulty of training Recurrent Neural Networks”. In : *ICML’2013*. 2013.

[267] Yoshua BENGIO, Grégoire MESNIL, Yann DAUPHIN et Salah RIFAI. “Better Mixing via Deep Representations”. In : *ICML’2013*. 2013.

[268] Heng LUO, Pierre Luc CARRIER, Aaron COURVILLE et Yoshua BENGIO. “Texture modeling with convolutional spike-and-slab RBMs and Deep Extensions”. In : *AISTATS’2013*. 2013.

[269] Grégoire MESNIL, Salah RIFAI, Antoine BORDES, Xavier GLOROT, Yoshua BENGIO et Pascal VINCENT. “Unsupervised and Transfer Learning under Uncertainty: from Object Detections to Scene Categorization”. In : *ICPRAM*. 2013.

[270] Salah RIFAI, Yoshua BENGIO, Yann DAUPHIN et Pascal VINCENT. “A Generative Process for Sampling Contractive Auto-Encoders”. In : *Proceedings of the Twenty-nine International Conference on Machine Learning (ICML’12)*. Edinburgh, Scotland, U.K. : ACM, 2012.

[271] Nicolas BOULANGER-LEWANDOWSKI, Yoshua BENGIO et Pascal VINCENT. “Discriminative Non-negative Matrix Factorization for Multiple Pitch Estimation”. In : *ISMIR*. 2012.

[272] Salah RIFAI, Yoshua BENGIO, Aaron COURVILLE, Pascal VINCENT et Mehdi MIRZA. “Disentangling factors of variation for facial expression recognition”. In : *European Conference on Computer Vision*. 2012.

[273] Antoine BORDES, Xavier GLOROT, Jason WESTON et Yoshua BENGIO. “Joint Learning of Words and Meaning Representations for Open-Text Semantic Parsing”. In : *AISTATS’2012* (2012).

[274] Ian J. GOODFELLOW, Aaron COURVILLE et Yoshua BENGIO. “Large-Scale Feature Learning With Spike-and-Slab Sparse Coding”. In : *Proceedings of the Twenty-nine International Conference on Machine Learning (ICML’12)*. ACM, 2012.

[275] N. BOULANGER-LEWANDOWSKI, Y. BENGIO et P. VINCENT. “Modeling Temporal Dependencies in High-Dimensional Sequences: Application to Polyphonic Music Generation and Transcription”. In : *ICML’2012*. 2012.

[276] Salah RIFAI, Yann DAUPHIN, Pascal VINCENT, Yoshua BENGIO et Xavier MULLER. “The Manifold Tangent Classifier”. In : *NIPS’2011*. 2011.

[277] James BERGSTRA, Rémy BARDENET, Yoshua BENGIO et Balázs KÉGL. “Algorithms for Hyper-Parameter Optimization”. In : *NIPS’2011*. 2011.

[278] G. DESJARDINS, A. COURVILLE et Y. BENGIO. “On Tracking The Partition Function”. In : *NIPS’2011*. 2011.

[279] Olivier DELALLEAU et Yoshua BENGIO. “Shallow vs. Deep Sum-Product Networks”. In : *NIPS*. 2011.

[280] Yoshua BENGIO et Olivier DELALLEAU. “On the Expressive Power of Deep Architectures”. In : *ALT’2011*. 2011.

[281] Yoshua BENGIO. “Deep Learning of Representations for Unsupervised and Transfer Learning”. In : *Journal of Machine Learning Research W&CP: Proc. Unsupervised and Transfer Learning*. 2011.

[282] Grégoire MESNIL, Yann DAUPHIN, Xavier GLOROT, Salah RIFAI, Yoshua BENGIO, Ian GOODFELLOW, Erick LAVOIE, Xavier MULLER, Guillaume DESJARDINS, David WARDE-FARLEY, Pascal VINCENT, Aaron COURVILLE et James BERGSTRA. “Unsupervised and Transfer Learning Challenge: a Deep Learning approach”. In : *Journal of Machine Learning Research W&CP: Proc. Unsupervised and Transfer Learning*. T. 7. 2011.

[283] Salah RIFAI, Grégoire MESNIL, Pascal VINCENT, Xavier MULLER, Yoshua BENGIO, Yann DAUPHIN et Xavier GLOROT. “Higher Order Contractive Auto-Encoder”. In : *ECML PKDD*. 2011.

[284] Antoine BORDES, Jason WESTON, Ronan COLLOBERT et Yoshua BENGIO. “Learning Structured Embeddings of Knowledge Bases”. In : *AAAI 2011*. 2011.

[285] Aaron COURVILLE, James BERGSTRA et Yoshua BENGIO. “Unsupervised Models of Images by Spike-and-Slab RBMs”. In : *ICML’2011*. 2011.

[286] Yann DAUPHIN, Xavier GLOROT et Yoshua BENGIO. “Large-Scale Learning of Embeddings with Reconstruction Sampling”. In : *Proceedings of theTwenty-eight International Conference on Machine Learning (ICML’11)*. Juin 2011.

[287] Salah RIFAI, Pascal VINCENT, Xavier MULLER, Xavier GLOROT et Yoshua BENGIO. “Contractive Auto-Encoders: Explicit invariance during feature extraction”. In : *ICML’2011*. 2011.

[288] Philippe HAMEL, Simon LEMIEUX, Yoshua BENGIO et Douglas ECK. “Temporal pooling and multiscale learning for automatic annotation and ranking of music audio”. In : *ISMIR*. 2011.

[289] Xavier GLOROT, Antoine BORDES et Yoshua BENGIO. “Domain Adaptation for Large-Scale Sentiment Classification: A Deep Learning Approach”. In : *ICML’2011*. 2011.

[290] X. GLOROT, A. BORDES et Y. BENGIO. “Deep Sparse Rectifier Neural Networks”. In : *AISTATS’2011*. 2011.

[291] Aaron COURVILLE, James BERGSTRA et Yoshua BENGIO. “A Spike and Slab Restricted Boltzmann Machine”. In : *Journal of Machine Learning Research W&CP: Proc. AISTATS’2011*. T. 15. 2011.

[292] Yoshua BENGIO et al. “Deep Learners Benefit More from Out-of-Distribution Examples”. In : *Journal of Machine Learning Research W&CP: Proc. AISTATS’2011*. 2011.

[293] James BERGSTRA, Olivier BREULEUX, Frédéric BASTIEN, Pascal LAMBLIN, Razvan PASCANU, Guillaume DESJARDINS, Joseph TURIAN, David WARDE-FARLEY et Yoshua BENGIO. “Theano: a CPU and GPU Math Expression Compiler”. In : *Proc. SciPy*. 2010.

[294] Michael MANDEL, Douglas ECK et Yoshua BENGIO. “Learning tags that vary within a song”. In : *In Proceedings of the 11th International Conference on Music Information Retrieval (ISMIR)*. Août 2010, p. 399–404.

[295] Joseph TURIAN, Lev RATINOV et Yoshua BENGIO. “Word representations: A simple and general method for semi-supervised learning”. In : *ACL’2010, ACL Test-of-Time Award 2020*. 2010, p. 384–394.

[296] Dumitru ERHAN, Aaron COURVILLE, Yoshua BENGIO et Pascal VINCENT. "Why Does Unsupervised Pre-training Help Deep Learning?" In : *Journal of Machine Learning Research W&CP: Proc. AISTATS'2010*. T. 9. 2010, p. 201–208.

[297] Xavier GLOROT et Yoshua BENGIO. "Understanding the difficulty of training deep feedforward neural networks". In : *AISTATS'2010*. 2010.

[298] Guillaume DESJARDINS, Aaron COURVILLE, Yoshua BENGIO, Pascal VINCENT et Olivier DELALLEAU. "Tempered Markov Chain Monte Carlo for training of Restricted Boltzmann Machine". In : *AISTATS*. 2010, p. 145–152.

[299] James BERGSTRA et Yoshua BENGIO. "Slow, Decorrelated Features for Pretraining Complex Cell-like Networks". In : *NIPS'2009*. Déc. 2009.

[300] Aaron COURVILLE, Douglas ECK et Yoshua BENGIO. "An Infinite Factor Model Hierarchy Via a Noisy-Or Mechanism". In : *Neural Information Processing Systems Conference (NIPS) 22*. 2009, p. 405–413.

[301] Yoshua BENGIO, Jerome LOURADOUR, Ronan COLLOBERT et Jason WESTON. "Curriculum Learning". In : *ICML'09*. 2009.

[302] Dumitru ERHAN, Pierre-Antoine MANZAGOL, Yoshua BENGIO, Samy BENGIO et Pascal VINCENT. "The Difficulty of Training Deep Architectures and the effect of Unsupervised Pre-Training". In : *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics (AISTATS 2009)*. Clearwater (Florida), USA, avr. 2009, p. 153–160.

[303] Joseph TURIAN, James BERGSTRA et Yoshua BENGIO. "Quadratic Features and Deep Architectures for Chunking". In : *Proc. NAACL-HLT'2009*. 2009, p. 245–248.

[304] Pascal VINCENT, Hugo LAROCHELLE, Yoshua BENGIO et Pierre-Antoine MANZAGOL. "Extracting and Composing Robust Features with Denoising Autoencoders". In : *ICML 2008*. 2008.

[305] Hugo LAROCHELLE et Yoshua BENGIO. "Classification using Discriminative Restricted Boltzmann Machines". In : *ICML'2008*. Helsinki, Finland, 2008, p. 536–543.

[306] Nicolas LE ROUX, Pierre-Antoine MANZAGOL et Yoshua BENGIO. "Topmoumoute online natural gradient algorithm". In : *Advances in Neural Information Processing Systems 20 (NIPS'07)*. Sous la dir. de J.C. PLATT, D. KOLLER, Y. SINGER et S. ROWEIS. Cambridge, MA : MIT Press, 2008, p. 849–856.

[307] Nicolas LE ROUX, Yoshua BENGIO, Pascal LAMBLIN, Marc JOLIVEAU et Balázs KÉGL. "Learning the 2-D Topology of Images". In : *Advances in Neural Information Processing Systems 20 (NIPS'07)*. Sous la dir. de John PLATT, D. KOLLER, Yoram SINGER et Sam T. ROWEIS. Cambridge, MA : MIT Press, 2008, p. 841–848.

[308] Nicolas CHAPADOS et Yoshua BENGIO. "Augmented Functional Time Series Representation and Forecasting with Gaussian Processes". In : *Advances in Neural Information Processing Systems 20 (NIPS'07)*. Sous la dir. de John PLATT, D. KOLLER, Yoram SINGER et Sam T. ROWEIS. Cambridge, MA : MIT Press, 2008, p. 265–272.

[309] Hugo LAROCHELLE, Dumitru ERHAN et Yoshua BENGIO. "Zero-data Learning of New Tasks". In : *AAAI Conference on Artificial Intelligence*. 2008.

[310] Hugo LAROCHELLE, Dumitru ERHAN, Aaron COURVILLE, James BERGSTRA et Yoshua BENGIO. "An Empirical Evaluation of Deep Architectures on Problems with Many Factors of Variation". In : *ICML'2007*. Corvallis, OR, 2007, p. 473–480.

[311] Yoshua BENGIO, Pascal LAMBLIN, Dan POPOVICI et Hugo LAROCHELLE. "Greedy Layer-Wise Training of Deep Networks". In : *NIPS'2006*. Sous la dir. de Bernhard SCHÖLKOPF, John PLATT et Thomas HOFFMAN. MIT Press, 2007, p. 153–160.

[312] Nicolas LE ROUX et Yoshua BENGIO. “Continuous Neural Networks”. In : *Proceedings of the Eleventh International Conference on Artificial Intelligence and Statistics (AISTATS’07)*. San Juan, Porto Rico : Omnipress, mar. 2007.

[313] Nicolas CHAPADOS et Yoshua BENGIO. “Forecasting Commodity Contract Spreads with Gaussian Process”. In : *13th International Conference on Computing in Economics and Finance*. Juin 2007.

[314] Julie CARREAU et Yoshua BENGIO. “A Hybrid Pareto Model for Conditional Density Estimation of Asymmetric Fat-Tail Data”. In : *AISTATS’2007*. Puerto Rico, 2007, p. 21–24.

[315] Yoshua BENGIO, Nicolas LE ROUX, Pascal VINCENT, Olivier DELALLEAU et Patrice MARCOTTE. “Convex Neural Networks”. In : *Advances in Neural Information Processing Systems 18 (NIPS’05)*. Sous la dir. d’Y. WEISS, B. SCHÖLKOPF et J. PLATT. Cambridge, MA : MIT Press, 2006, p. 123–130.

[316] Yoshua BENGIO, Olivier DELALLEAU et Nicolas LE ROUX. “The Curse of Highly Variable Functions for Local Kernel Machines”. In : *NIPS’2005*. 2006.

[317] Yoshua BENGIO, Hugo LAROCHELLE et Pascal VINCENT. “Non-Local Manifold Parzen Windows”. In : *Advances in Neural Information Processing Systems 18 (NIPS’05)*. Sous la dir. d’Y. WEISS, B. SCHÖLKOPF et J. PLATT. MIT Press, 2006, p. 115–122.

[318] Nicolas CHAPADOS et Yoshua BENGIO. “The K Best-Paths Approach to Approximate Dynamic Programming with Application to Portfolio Optimization”. In : *AI06*. 2006, p. 491–502.

[319] Yves GRANDVALET et Yoshua BENGIO. “Semi-supervised Learning by Entropy Minimization”. In : *Advances in Neural Information Processing Systems 17 (NIPS’04)*. Sous la dir. de L.K. SAUL, Y. WEISS et L. BOTTOU. Cambridge, MA : MIT Press, déc. 2005.

[320] F. RIVEST, Y. BENGIO et J. KALASKA. “Brain Inspired Reinforcement Learning”. In : *Advances in Neural Information Processing Systems 17 (NIPS’04)*. Sous la dir. de L.K. SAUL, Y. WEISS et L. BOTTOU. Cambridge, MA : MIT Press, Cambridge, 2005, p. 1129–1136.

[321] Yoshua BENGIO et Martin MONPERRUS. “Non-Local Manifold Tangent Learning”. In : *Advances in Neural Information Processing Systems 17 (NIPS’04)*. Sous la dir. de L.K. SAUL, Y. WEISS et L. BOTTOU. MIT Press, 2005, p. 129–136.

[322] Olivier DELALLEAU, Yoshua BENGIO et Nicolas LE ROUX. “Efficient Non-Parametric Function Induction in Semi-Supervised Learning”. In : *Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics (AISTATS’05)*. Sous la dir. de Robert G. COWELL et Zoubin GHAHRAMANI. Savannah Hotel, Barbados : Society for Artificial Intelligence et Statistics, jan. 2005, p. 96–103.

[323] Marie OUIMET et Yoshua BENGIO. “Greedy Spectral Embedding”. In : *AISTATS’2005*. Savannah Hotel, Barbados, 2005, p. 253–260.

[324] Frédéric MORIN et Yoshua BENGIO. “Hierarchical Probabilistic Neural Network Language Model”. In : *AISTATS’2005*. 2005, p. 246–252.

[325] I. BHATTACHARYA, L. GETOOR et Y. BENGIO. “Unsupervised Sense Disambiguation Using Bilingual Probabilistic Models”. In : *Conference of the Association for Computational Linguistics (ACL’04)*. 2004.

[326] Yoshua BENGIO, Jean-Francois PAIEMENT, Pascal VINCENT, Olivier DELALLEAU, Nicolas LE ROUX et Marie OUIMET. “Out-of-Sample Extensions for LLE, Isomap, MDS, Eigenmaps, and Spectral Clustering”. In : *Advances in Neural Information Processing Systems 16 (NIPS’03)*. Sous la dir. de S. THRUN, L. SAUL et B. SCHÖLKOPF. MIT Press, 2004.

[327] Y. BENGIO et Y. GRANDVALET. “No Unbiased Estimator of the Variance of K-Fold Cross-Validation”. In : *Advances in Neural Information Processing Systems 16 (NIPS’03)*. Sous la dir. de S. THRUN, L. SAUL et B. SCHÖLKOPF. Cambridge, MA : MIT Press, Cambridge, 2004.

[328] N. BOUFADEN, Y. BENGIO et G. LAPALME. “Approche statistique pour le repérage de mots informatifs dans les textes oraux”. In : TALN’2004*, Traitement Automatique du Langage Naturel.* 2004.

[329] Yoshua BENGIO et Jean-Sébastien SÉNÉCAL. “Quick Training of Probabilistic Neural Nets by Importance Sampling”. In : *Proceedings of the conference on Artificial Intelligence and Statistics (AISTATS)*. 2003.

[330] Pascal VINCENT et Yoshua BENGIO. “Manifold Parzen Windows”. In : *Advances in Neural Information Processing Systems 15 (NIPS’02)*. Sous la dir. de S. BECKER et Sebastian THRUN. Cambridge, MA : MIT Press, 2003, p. 825–832.

[331] R. COLLOBERT, Y. BENGIO et S. BENGIO. “Scaling Large Learning Problems with Hard Parallel Mixtures”. In : *Pattern Recognition with Support Vector Machines*. Sous la dir. de S.W. LEE et A. VERRI. T. 2388 of Lecture Notes in Computer Science. Springer-Verlag, 2002, p. 8–23.

[332] Y. BENGIO, I. TAKEUCHI et K. KANAMORI. “The Challenge of Non-Linear Regression on Large Datasets with Asymmetric Heavy Tails”. In : *Proceedings of 2002 Joint Statistical Meetings*. American Statistical Association publ., 2002, p. 193–205.

[333] Y. BENGIO et N. CHAPADOS. “Metric-based Model Selection for Time-Series Forecasting”. In : *Proc. of 2002 IEEE International Workshop on Neural Networks for Signal Processing*. Martigny, Suisse : IEEE Press, 2002, p. 13–24.

[334] P. VINCENT et Y. BENGIO. “K-Local Hyperplane and Convex Distance Nearest Neighbor Algorithms”. In : *Advances in Neural Information Processing Systems 14 (NIPS’01)*. Sous la dir. de T.G. DIETTERICH, S. BECKER et Z. GHAHRAMANI. Cambridge, MA : MIT Press, 2002, p. 985–992.

[335] R. COLLOBERT, S. BENGIO et Y. BENGIO. “A Parallel Mixture of SVMs for Very Large Scale Problems”. In : *Advances in Neural Information Processing Systems 14 (NIPS’01)*. Sous la dir. de T.G. DIETTERICH, S. BECKER et Z. GHAHRAMANI. 2002, p. 633–640.

[336] N. CHAPADOS, Y. BENGIO, P. VINCENT, J. GHOSN, C. DUGAS, I. TAKEUCHI et L. MENG. “Estimating Car Insurance Premia: a Case Study in High-Dimensional Data Inference”. In : *Advances in Neural Information Processing Systems 14 (NIPS’01)*. Sous la dir. de T.G. DIETTERICH, S. BECKER et Z. GHAHRAMANI. Cambridge, MA : MIT Press, 2002, p. 1369–1376.

[337] N. BOUFADEN, Lapalme G. et Y. BENGIO. “Topic segmentation: First Stage of Dialogue-Based Information extraction Process”. In : *Proceedings of the Natural Language Pacific Rim Symposium, NLPRS-01*. 2001.

[338] Yoshua BENGIO, Réjean DUCHARME et Pascal VINCENT. “A Neural Probabilistic Language Model”. In : *Advances in Neural Information Processing Systems 13 (NIPS’00)*. Sous la dir. de T.K. LEEN, T.G. DIETTERICH et V. TRESP. MIT Press, 2001, p. 932–938.

[339] C. DUGAS, Y. BENGIO, F. BÉLISLE et C. NADEAU. “Incorporating Second-Order Functional Knowledge for Better Option Pricing”. In : *Advances in Neural Information Processing Systems 13 (NIPS’00)*. Sous la dir. de T.K. LEEN, T.G. DIETTERICH et V. TRESP. MIT Press, 2001, p. 472–478.

[340] Y. BENGIO. “Probabilistic neural network models for sequential data”. In : *International Joint Conference on Neural Networks (IJCNN)*. T. 5. 2000, p. 79–84.

[341] Yoshua BENGIO. “Continuous Optimization of Hyper-Parameters”. In : *International Joint Conference on Neural Networks (IJCNN)*. T. V. 2000, p. 305–310.

[342] J. GHOSN et Y. BENGIO. “Bias Learning, Knowledge Sharing”. In : *International Joint Conference on Neural Networks (IJCNN)*. T. I. 2000, p. 9–14.

[343] Pascal VINCENT et Yoshua BENGIO. “A Neural Support Vector Network Architecture with Adaptive Kernels”. In : *International Joint Conference on Neural Networks (IJCNN)*. T. 5. 2000, p. 5187–5192.

[344] N. CHAPADOS et Y. BENGIO. “VaR-based Asset Allocation using Neural Networks”. In : *Computational Finance 2000*. 2000.

[345] F. GINGRAS, Y. BENGIO et C. NADEAU. “On Out-of-Sample Statistics for Time-Series”. In : *Computational Finance 2000*. London, U.K. : 2000.

[346] Yoshua BENGIO et Samy BENGIO. “Modeling High-Dimensional Discrete Data with Multi-Layer Neural Networks”. In : *NIPS’1999*. MIT Press, 2000, p. 400–406.

[347] Claude NADEAU et Yoshua BENGIO. “Inference for the Generalization Error”. In : *Advances in Neural Information Processing Systems 12 (NIPS’99)*. Sous la dir. de S.A. SOLLA, T.K. LEEN et K-R. MÜLLER. MIT Press, 2000, p. 307–313.

[348] S. PIGEON et Y. BENGIO. “Binary Pseudowavelets and Application to Bilevel Image Processing”. In : *Proceedings of the Data Compression Conference, DCC’1999*. 1999.

[349] Y. BENGIO, S. BENGIO, J. F. ISABELLE et Y. SINGER. “Shared Context Probabilistic Transducers”. In : *Advances in Neural Information Processing Systems 10 (NIPS’97)*. Sous la dir. de M.I. JORDAN, M.J. KEARNS et S.A. SOLLA. MIT Press, 1998, p. 409–415.

[350] M. BONNEVILLE, J. MEUNIER, Y. BENGIO et J.P. SOUCY. “Support Vector Machines for Improving the classification of Brain Pet Images”. In : *SPIE Medical Imaging 1998*. San Diego, 1998.

[351] Holger SCHWENK et Yoshua BENGIO. “Training Methods for Adaptive Boosting of Neural Networks”. In : *Advances in Neural Information Processing Systems 10 (NIPS’97)*. Sous la dir. de M.I. JORDAN, M.J. KEARNS et S.A. SOLLA. MIT Press, 1998, p. 647–653.

[352] P. HAFFNER, L. BOTTOU, P. G. HOWARD, P. SIMARD, Y. BENGIO et Y. Le CUN. “Browsing through High Quality Document Images with DjVu”. In : *Proceedings of the Advances in Digital Libraries Conference (ADL’98)*. Washington, DC, USA : IEEE Computer Society, 1998, p. 309. ISBN : 0-8186-8464-X.

[353] L. BOTTOU, P. G. HOWARD et Y. BENGIO. “The Z-Coder Adaptive Binary Coder”. In : *Proceedings of the Conference on Data Compression (DCC’98)*. Washington, DC, USA : IEEE Computer Society, 1998, p. 13.

[354] Steven PIGEON et Yoshua BENGIO. “A Memory-Efficient Adaptive Huffman Coding Algorithm for Very Large Sets of Symbols”. In : *Proceedings of the Conference on Data Compression (DCC’98)*. 1998, p. 568.

[355] Yann LECUN, L. BOTTOU et Y. BENGIO. “Reading checks with multilayer graph transformer networks”. In : *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP’97)*. T. 1. 1997, p. 151–154.

[356] Mazin RAHIM, Yoshua BENGIO et Yann LECUN. “Discriminative Feature And Model Design For Automatic Speech Recognition”. In : *In Proc. of Eurospeech*. 1997, p. 75–78.

[357] Léon BOTTOU, Yoshua BENGIO et Yann LECUN. “Global Training of Document Processing Systems using Graph Transformer Networks”. In : *Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR’97)*. Puerto Rico : IEEE, 1997, p. 490–494.

[358] Holger SCHWENK et Yoshua BENGIO. “AdaBoosting Neural Networks: Application to on-line Character Recognition”. In : *International Conference on Artificial Neural Networks*. Springer Verlag, 1997, p. 967–972.

[359] Joumana GHOSN et Yoshua BENGIO. “Multi-Task Learning for Stock Selection”. In : *NIPS’1996*. MIT Press, Cambridge, MA, 1997, p. 946–952.

[360] Yoshua BENGIO. “Training A Neural Network with a Financial Criterion Rather then a Prediction Criterion”. In : *Proceedings of the Fourth International Conference on Neural Networks in the Capital Markets (NNCM-96)*. Sous la dir. d’A.S. WEIGEND, Y.S ABU-MOSTAFA et A.-P.N. REGENES. World Scientific, 1997, p. 433–443.

[361] Samy BENGIO et Yoshua BENGIO. “An EM Algorithm for Asynchronous Input/Output Hidden Markov Models”. In : *International Conference On Neural Information Processing*. Sous la dir. de L. XU. Hong-Kong, 1996, p. 328–334.

[362] Yoshua BENGIO et François GINGRAS. “Recurrent Neural Networks for Missing or Asynchronous Data”. In : *NIPS’1995*. 1996, p. 395–401.

[363] Salah EL HIHI et Yoshua BENGIO. “Hierarchical Recurrent Neural Networks for Long-Term Dependencies”. In : *Advances in Neural Information Processing Systems 8 (NIPS’95)*. Sous la dir. de D. S. TOURETZKY, M. MOZER et M.E. HASSELMO. MIT Press, 1996.

[364] Yoshua BENGIO et Paolo FRASCONI. “An Input/Output HMM Architecture”. In : *NIPS’1994*. 1995, p. 427–434.

[365] Yoshua BENGIO et Paolo FRASCONI. “Diffusion of Credit in Markovian Models”. In : *NIPS’1994*. 1995, p. 553–560.

[366] Leon BOTTOU et Yoshua BENGIO. “Convergence Properties of the K-Means Algorithm”. In : *Advances in Neural Information Processing Systems 7 (NIPS’94)*. Sous la dir. de Gerald TESAURO, D. S. TOURETZKY et T. K. LEEN. MIT Press, 1995, p. 585–592.

[367] S. BENGIO, Y. BENGIO et J. CLOUTIER. “Use of genetic programming for the search of a new learning rule for neural networks”. In : *Proceedings of the First IEEE Conference on Evolutionary Computation*. 1994, 324–327 vol.1.

[368] Paolo FRASCONI et Yoshua BENGIO. “An EM Approach to Grammatical Inference: Input/Output HMMs”. In : *International Conference on Pattern Recognition (ICPR’94)*. Jerusalem 1994, 1994, p. 289–294.

[369] Yann LECUN et Yoshua BENGIO. “Word-Level Training of a Handritten Word Recognizer based on Convolutional Neural Networks”. In : *International Conference on Pattern Recognition (ICPR’94)*. Sous la dir. d’IEEE. Jerusalem 1994, 1994.

[370] Yoshua BENGIO et Yann LECUN. “Word Normalization For On-Line Handwritten Word Recognition”. In : *International Conference on Pattern Recognition (ICPR’94)*. 1994, p. 409–413.

[371] Yoshua BENGIO, Yann LECUN et Donnie HENDERSON. “Globally Trained Handwritten Word Recognizer using Spatial Representation, Space Displacement Neural Networks and Hidden Markov Models”. In : *Advances in Neural Information Processing Systems 6 (NIPS’93)*. Sous la dir. de J. D. COWAN, Gerald TESAURO et J. ALSPECTOR. MIT Press, 1994, p. 937–944.

[372] Yoshua BENGIO et Paolo FRASCONI. “Credit Assignment through Time: Alternatives to Backpropagation”. In : *NIPS’1993*. 1994, p. 75–82.

[373] Y LECUN, Y. BENGIO, D. HENDERSON et A. WEISBUCH. “On-Line handwriting recognition with neural networks: spatial representation versus temporal representation”. In : *Proceedings of the International Conference on Handwriting and Drawing*. Ecole Nationale Superieure des Telecommunications, 1993.

[374] Yoshua BENGIO, Paolo FRASCONI, Marco GORI et G. SODA. “Recurrent Neural Networks for Adaptive Temporal Processing”. In : *Proc. of the 6th Italian Workshop on Neural Networks, WIRN-93*. Sous la dir. d’E. CAIANELLO. Vietri, Italy : World Scientific Publ., 1993, p. 1183–1195.

[375] S. BENGIO, Y. BENGIO, J. CLOUTIER et J. GECSEI. “Generalization of a Parametric Learning Rule”. In : *Proceedings of the International Conference on Artificial Neural Networks 1993*. Sous la dir. de S. GIELEN et B. KAPPEN. Amsterdam, The Netherlands : Springer-Verlag, 1993, p. 502–502.

[376] Y. BENGIO, P. FRASCONI et P. SIMARD. “The problem of learning long-term dependencies in recurrent networks”. In : *IEEE International Conference on Neural Networks*. San Francisco : IEEE Press, 1993, p. 1183–1195.

[377] Yoshua BENGIO, Renato DE MORI, Giovanni FLAMMIA et Ralf KOMPE. “Neural Network - Gaussian Mixture Hybrid for Speech Recognition or Density Estimation”. In : *Advances in Neural Information Processing Systems 4 (NIPS’91)*. Sous la dir. de J. E. MOODY, S. J. HANSON et R. P. LIPMANN. Denver, CO : Morgan Kaufmann, 1992, p. 175–182.

[378] Samy BENGIO, Yoshua BENGIO, Jocelyn CLOUTIER et Jan GECSEI. “Aspects théoriques de l’optimisation d’une règle d’apprentissage”. In : *Actes de la conférence Neuro-Nımes 1992*. Nımes, France, 1992.

[379] Yoshua BENGIO, Samy BENGIO, Jocelyn CLOUTIER et Jan GECSEI. “On the Optimization of a Synaptic Learning Rule”. In : *in Conference on Optimality in Biological and Artificial Networks*. 1992.

[380] Y. BENGIO, R. DE MORI, G. FLAMMIA et R. KOMPE. “Global Optimization of a Neural Network - Hidden Markov Model Hybrid”. In : *International Joint Conference on Neural Networks (IJCNN)*. T. 2. 1991, p. 789–794.

[381] Yoshua BENGIO, Renato DE MORI, Giovanni FLAMMIA et Ralf KOMPE. “A comparative study of hybrid acoustic phonetic decoders based on artificial neural networks”. In : *Proceedings of EuroSpeech’91*. 1991.

[382] Yoshua BENGIO, Samy BENGIO, Jocelyn CLOUTIER et Jan GECSEI. “Learning a Synaptic Learning Rule”. In : *IJCNN’1991*. Seattle, WA, 1991, II–A969.

[383] Yoshua BENGIO, Renato De MORI, Giovanni FLAMMIA et Ralf KOMPE. “A Comparative Study On Hybrid Acoustic Phonetic Decoders Based On Artificial Neural Networks”. In : *Proceeding of EuroSpeech*. Genova, Italy, 1991.

[384] Yoshua BENGIO, Renato DE MORI et Marco GORI. “Experiments on automatic speech recognition using BPS”. In : *Parallel Architectures and Neural Networks*. Sous la dir. d’E. CAIANELLO. World Scientific Publ., 1990, p. 223–232.

[385] Yoshua BENGIO, Regis CARDIN, Renato DE MORI et Yves NORMANDIN. “A Hybrid Coder for Hidden Markov Models Using a Recurrent Neural Network”. In : *International Conference on Acoustics, Speech and Signal Processing*. Albuquerque, NM, 1990, p. 537–540.

[386] Yoshua BENGIO, Yannick POULIOT, Samy BENGIO et Patrick AGIN. “A neural network to detect homologies in proteins”. In : *Advances in Neural Information Processing Systems 2 (NIPS’89)*. Sous la dir. de D. S. TOURETZKY. Denver, CO : Morgan Kaufmann, 1990, p. 423–430.

[387] Yoshua BENGIO, Regis CARDIN et Renato DE MORI. “Speaker Independent Speech Recognition with Neural Networks and Speech Knowledge”. In : *Advances in Neural Information Processing Systems 2 (NIPS’89)*. Sous la dir. de D. S. TOURETZKY. Denver, CO : Morgan Kaufmann, 1990, p. 218–225.

[388] Y. BENGIO, P. COSI, R. CARDIN et R. De MORI. “Use of multi-layered networks for coding speech with phonetic features”. In : *Advances in Neural Information Processing Systems 1 (NIPS’88)*. Sous la dir. de D.S. TOURETZKY. Denver, CO : Morgan Kaufmann, San Mateo, 1989, p. 224–231.

[389] Renato DE MORI, Yoshua BENGIO et Piero COSI. “On the generalization capability of multilayered networks in the extraction of speech properties”. In : *IJCAI’1989*. Detroit : IEEE, 1989, p. 1531–1536.

[390] M. GORI, Y. BENGIO et R. DE MORI. “BPS: A Learning Algorithm for Capturing the Dynamical Nature of Speech”. In : *International Joint Conference on Neural Networks (IJCNN)*. Washington D.C. : IEEE, New York, 1989, p. 643–644.

[391] Yoshua BENGIO, Regis CARDIN, Piero COSI et Renato DE MORI. “Speech coding with multi-layer networks”. In : *International Conference on Acoustics, Speech and Signal Processing*. Glasgow, Scotland, 1989, p. 164–167.

[392] Yoshua BENGIO et Renato DE MORI. “Use of neural networks for the recognition of place of articulation”. In : *ICASSP’1988*. New-York, NY, 1988, p. 103–106.

Books and Book Chapters

[393] Ian J. GOODFELLOW, Yoshua BENGIO et Aaron COURVILLE. *Deep Learning*. MIT Press, 2016.

[394] Yoshua BENGIO. “Evolving Culture vs Local Minima”. In : *Growing Adaptive Machines: Integrating Development and Learning in Artificial Neural Networks*. also as ArXiv 1203.2990v1. Springer-Verlag, mar. 2013, T. Kowaliw, N. Bredeche & R. Doursat, eds.

[395] Yoshua BENGIO et Aaron COURVILLE. “Deep Learning of Representations”. In : *Handbook on Neural Information Processing*. T. 49. Springer: Berlin Heidelberg, 2013.

[396] Yoshua BENGIO. *Learning deep architectures for AI*. Now Publishers, 2009.

[397] Yoshua BENGIO et Yann LECUN. “Scaling Learning Algorithms towards AI”. In : *Large Scale Kernel Machines*. Sous la dir. de L. BOTTOU, O. CHAPELLE, D. DECOSTE et J. WESTON. MIT Press, 2007.

[398] Yves GRANDVALET et Yoshua BENGIO. “Entropy Regularization”. In : *Semi-Supervised Learning*. Sous la dir. d’Olivier CHAPELLE, Bernhard SCHÖLKOPF et Alexander ZIEN. MIT Press, 2006, p. 151–168.

[399] Yoshua BENGIO, Olivier DELALLEAU et Nicolas LE ROUX. “Label Propagation and Quadratic Criterion”. In : *Semi-Supervised Learning*. Sous la dir. d’Olivier CHAPELLE, Bernhard SCHÖLKOPF et Alexander ZIEN. MIT Press, 2006, p. 193–216.

[400] Olivier DELALLEAU, Yoshua BENGIO et Nicolas LE ROUX. “Large-Scale Algorithms”. In : *Semi-Supervised Learning*. Sous la dir. d’Olivier CHAPELLE, Bernhard SCHÖLKOPF et Alexander ZIEN. MIT Press, 2006, p. 333–341.

[401] Yoshua BENGIO, Olivier DELALLEAU, Nicolas LE ROUX, Jean-François PAIEMENT, Pascal VINCENT et Marie OUIMET. “Spectral Dimensionality Reduction”. In : *Feature Extraction, Foundations and Applications*. Sous la dir. d’Isabelle GUYON, Steve GUNN, Masoud NIKRAVESH et Lofti ZADEH. Springer, 2006.

[402] Y. BENGIO et Y. GRANDVALET. “Bias in Estimating the Variance of K-Fold Cross-Validation”. In : *Statistical Modeling and Analysis for Complex Data Problem*. Sous la dir. de P. DUCHESNE et B. REMILLARD. Kluwer : Lawrence Erlbaum, 2004, p. 75–95.

[403] C. DUGAS, Y. BENGIO, N. CHAPADOS, P. VINCENT, G. DENONCOURT et C. FOURNIER. "Statistical Learning Algorithms Applied to Automobile Insurance Ratemaking". In : *Intelligent and Other Computational Techniques in Insurance: Theory and Applications*. Sous la dir. de L. JAIN et A.F. SHAPIRO. World Scientific Publishing Company, 2004.

[404] E. TRENTIN, F. BRUGNARA, Y. BENGIO, C. FURLANELLO et R. De MORI. "Statistical and Neural Network Models for Speech Recognition". In : *Connectionist Approaches to Clinical Problems in Speech and Language*. Sous la dir. de R. DANILOFF. Lawrence Erlbaum, 2002, p. 213–264.

[405] Y. LECUN, L. BOTTOU, Y. BENGIO et P. HAFFNER. "Gradient-Based Learning Applied to Document Recognition". In : *Intelligent Signal Processing*. IEEE Press, 2001, p. 306–351.

[406] Juergen SCHMIDHUBER, Sepp HOCHREITER et Yoshua BENGIO. "Evaluating Benchmark Problems by Random Guessing". In : *Field Guide to Dynamical Recurrent Networks*. Sous la dir. de J. KOLEN et S. KREMER. IEEE Press, 2001.

[407] Sepp HOCHREITER, Fakultat F. INFORMATIK, Yoshua BENGIO, Paolo FRASCONI et Jurgen SCHMIDHUBER. "Gradient Flow in Recurrent Nets: the Difficulty of Learning Long-Term Dependencies". In : *Field Guide to Dynamical Recurrent Networks*. Sous la dir. de J. KOLEN et S. KREMER. IEEE Press, 2000.

[408] Yann LECUN, Patrick HAFFNER, Leon BOTTOU et Yoshua BENGIO. "Object Recognition with Gradient-Based Learning". In : *Shape, Contour and Grouping in Computer Vision*. Springer, 1999, p. 319–345.

[409] Yoshua BENGIO. *Neural Networks for Speech and Sequence Recognition*. London, UK : International Thompson Computer Press, 1996.

[410] Yann LECUN et Yoshua BENGIO. "Convolutional Networks for Images, Speech, and Time-Series". In : *The Handbook of Brain Theory and Neural Networks*. Sous la dir. de M. A. ARBIB. MIT Press, 1995, p. 255–257.

[411] Yann LECUN et Yoshua BENGIO. "Pattern Recognition and Neural Networks". In : *The Handbook of Brain Theory and Neural Networks*. Sous la dir. de M. A. ARBIB. MIT Press, 1995, p. 711–714.

[412] Yoshua BENGIO. "Radial Basis Functions for speech recognition". In : *Speech Recognition and Understanding: Recent Advances, Trends and Applications*. NATO Advanced Study Institute Series F: Computer et Systems Sciences, 1990, p. 293–298.

[413] Yoshua BENGIO et Renato DE MORI. "Speech coding with multilayer networks". In : *Neurocomputing: Algorithms, Architectures and Applications*. Sous la dir. de F. FOGELMAN SOULIE et J. HERAULT. NATO Advanced Study Institute Series F: Computer et Systems Sciences, 1990, p. 207–216.

[414] Renato DE MORI, Yoshua BENGIO et Piero COSI. "On the use of an ear model and multi-layer networks for automatic speech recognition". In : *Structural Pattern Analysis*. Sous la dir. de R. MOHR, T. PAVLIDIS et A. SANFELIN. World Scientific, 1990.

[415] Yoshua BENGIO et Renato DE MORI. "Connectionist models and their application to automatic speech recognition". In : *Artificial Neural Networks and Statistical Pattern Recognition: Old and New Connections*. Sous la dir. d'I. K. SETHI et A. K. JAIN. Elsevier, Machine Intelligence et Pattern Recognition Series, 1990, p. 175–192.

Patents

[416] Nicolas CHAPADOS, Nicolas GUIZARD, Mohammad HAVAEI et Yoshua BENGIO.

Method and system for processing a task with robustness to missing input information. US Patent App. 16/085,339. 2019.

[417] Yoshua BENGIO, Leon BOTTOU et Paul G. HOWARD. *Z-Coder : a fast adaptive binary arithmetic coder*. U.S. Patent 6,188,334, February 13, 2001, along with patents 6,225,925, 6,281,817, and 6,476,740. 2001.

[418] Yoshua BENGIO, Leon BOTTOU et Yann LECUN. *Module for constructing trainable modular network in which each module outputs and inputs data structured as a graph*. U.S. Patent 6,128,606, October 3. 2000.

[419] Yoshua BENGIO, Yann LECUN, Craig NOHL et Chris BURGES. *Visitor Registration System Using Automatic Handwriting Recognition*. Patent submitted in the U.S.A. in October 1994, submission number 1-16-18-1. 1994.