Mohammad Pezeshki

Montréal Institute for Learning Algorithms, Department of Computer Science and Operations Research
Université de Montréal, 2900 Boulevard Edouard-Montpetit, Montréal, QC H3T 1J4
Email: mohammad.pezeshki@umontreal.ca

RESEARCH INTERESTS

- Sequence Data Modeling
- Attention-based Neural Networks
- Recurrent Neural Networks
- Representation Learning

EDUCATION

- **Université de Montréal,**, Montréal, Canada
  Ph.D. Candidate under the supervision of Yoshua Bengio and Aaron Courville, Computer Science *Sep. 2016 _ Present*

- **Université de Montréal,**, Montréal, Canada
  M.Sc. under the supervision of Yoshua Bengio and Aaron Courville, Computer Science *Sep. 2014 _ June. 2016*
  GPA: **3.92** / 4.3

- **Amirkabir University of Technology** (Tehran Polytechnic), Tehran, Iran
  BS, Computer Engineering, *Sep. 2010 _ June 2014*
  GPA: **18.48** / 20

HONORS AND AWARDS

- Department of Computer Science and Operation Research (DIRO) M.Sc. and Ph.D. scholarship, University of Montreal, 2014 and 2016.
- M.Sc. and Ph.D. Tuition Exception scholarships (Bource C), University of Montreal, 2014 and 2016.
- Department of Computer Science and Operation Research Excellence Award, University of Montreal, Sept. 2015.
- Awarded as Outstanding Student in Amirkabir University of Technology, Tehran, Iran, 2011, 2012, 2013, 2014.
- Ranked 2nd in Cumulative GPA among 105 undergraduate students in Computer Engineering and IT Department, Amirkabir University of Technology, Tehran, Iran, 2014.
- Direct admission to Graduate Program at the Computer Engineering and IT Department of *Amirkabir University of Technology* and *Sharif University of Technology*, 2014.
- Ranked in the top 0.5% among all applicants for the Nationwide University Entrance Exam (Approximately 300000 applicants) Iran, 2010.
- Admitted to the National Organization for Development of Exceptional Talents (NODET) for high school studies, 2006.

PUBLICATIONS

- Krueger, D., Maharaj, T., Kramr, J., Pezeshki, M., Ballas, N., Ke, N. R., et al. & Pal, C. (2016). Zoneout: Regularizing RNNs by Randomly Preserving Hidden Activations. *arXiv preprint arXiv:1606.01305.*
- Al-Rfou, R., Alain, G., Almahairi, A., et al, Pezeshki, M., et al & Angermueller, C. (2016). Theano: A Python framework for fast computation of mathematical expressions. *arXiv eprints arXiv:1605.02688.*
- Pezeshki, M., Fan, L., Brakel, P., Courville, A., & Bengio, Y. (2016). Deconstructing the Ladder Network Architecture. *Proceedings of the 33rd International Conference on Machine Learning (ICML), 2016.*
- Zhang, Y., Pezeshki, M., Brakel, P., Zhang, S., Laurent, C., Bengio, Y., & Courville, A. (2016). Towards End-to-End Speech Recognition with Deep Convolutional Neural Networks. *INTERSPEECH 2016.*
- Pezeshki, M. (2015). Sequence Modeling using Gated Recurrent Neural Networks. *arXiv preprint arXiv:1501.00299.*
- Keyvanrad, M.A., Pezeshki, & M., Homayounpour, M.M. (2014). Deep Belief Networks for Image Denoising. *Workshop paper in International Conference on Learning Representations (ICLR 2014).*
- Pezeshki, M., Gholami, S., & Nickabadi, A. (2014). *"Distinction between features extracted using Deep Belief Networks". arXiv preprint arXiv:1312.6157.*
- Gohari, E., Mehri, S., Sadeghi, S., Pezeshki, M., Rashidi, F., Malakouti, S., & Khadivi, S. (2014). A Study on Prediction of User's Tendency Toward Purchases in Online Websites Based on Behavior Models. *6th Conference on Information and Knowledge Technology (IKT 2014).*

PROFESSIONAL EXPERIENCES

- **Research Intern at Twitter Inc.** ,
  Under the supervision of Hugo Larochelle *(July. 2016 _ Present)*
  Research subject to NDA.
- **Thesis based Ph.D. and M.Sc. Student and Research Assistant**
  Montréal Institute for Learning Algorithms (MILA), *(Sept. 2014 _ Present)*
  Under the supervision of Yoshua Bengio and Aaron Courville
  Artificial Neural Network Research with focus on time-series data modeling including text, speech, and video. Working specifically with Recurrent models such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), as well as feed forward models such as Convolutional and Residual Networks.
- **Research Assistant**
  Intelligent Multimedia Processing Lab, Human Language Technology Lab, *(Nov. 2012 _ Sept. 2014)*
  Under the supervision of Shahram Khadivi, MohammadMehdi Homayounpour and MohammadAli Keyvanrad
  Developing methods based on Restricted Boltzmann Machine (RBM) and Deep Belief Network (DBN) for denoising image data. Part of research was also focused on analyzing online shopping data using data mining algorithms including Decision trees and SVM classifiers.
- **Teaching Assistant**,
  (Fall 2012, Spring 2012, Fall 2013)
  Programming Languages, Data Structures, Advanced Programming

TOP ACADEMIC PROJECTS

- **Attention on Videos**
  Developed a method that extracts patches from frames of a video followed by the application of Recurrent Neural Network leading to classification after processing the whole sequence. This project is implemented using Blocks and Theano. [*Github link*]
- **End-to-end Speech Recognition**
  Implemented Connectionist Temporal Classification (CTC) on top of Recurrent Neural Networks with Long Short-Term Memory (LSTM) units. This project is implemented using Theano. [*Github link (only CTC code is public)*]
- **Character-level Language Modeling**
  Implemented a multi-layer network for Modeling a sequence of characters on datasets of PennTreebank and Wikipedia using recurrent models of Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM). This project is implemented using Theano and Blocks in python. [*Github link*]
- **Cats vs Dogs**
  Developed a Convolutional Neural Network armed with Batch Normalization for binary classification on 25000 images of cats and dogs. The code is implemented using Theano in python. [*Github link*]
- **LSTM in Blocks**
  Implemented the Long Short-Term Memory (LSTM) module that was added to the Blocks library on top of Theano. [*Github link*]
- **Information Retrieval**
  Implemented a static search engine using both traditional TF-IDF method and Semantic Hashing for content-based document retrieval. This project is in Java and Matlab.

TECHNICAL SKILLS

- **Programming Languages:**
  Expert in: Python, Matlab, Java
  Familiar with: C++
- **Tools and Frameworks:**
  Expert in: Theano, Torch, Blocks
  Familiar with: Tensorflow

REFERENCES

- **Prof. Yoshua Bengio** and **Prof. Aaron Courville**
  Montréal Institute for Learning Algorithms (MILA), Department of Computer Science and Operations Research, Université de Montréal.
  Emails: `bengioy@iro.umontreal.ca`, `aaron.courville@umontreal.ca`
- **Prof. Hugo Larochelle**
  Research Scientist at Twitter Inc., on leave Assistant Professor at Université de Sherbrooke.
  Email: `hlarochelle@twitter.com`
- **Professor Ahmad Nickabadi**
  Computer Engineering and IT Department, Amirkabir University of Technology.
  Email: `nickabadi@aut.ac.ir`