Aparna Balagopalan

Website: www.cs.toronto.edu/ aparna/
Google Scholar: Link

Email : aparna@cs.toronto.edu
Github: https://github.com/Aparna-B

Research Interests

Developing and understanding trustworthy machine learning and natural language processing techniques for healthcare applications: particularly using unstructured and noisy human-generated data, with core machine learning methodologies such as transfer learning and reinforcement learning.

Education

- **University of Toronto** — Toronto, Canada
  Ph.D. in Computer Science — *Sep 2020 - Present*
- **University of Toronto** — Toronto, Canada
  MSc. in Applied Computing; Advisor: Marzyeh Ghassemi — *Sep 2017 - Jan 2019*
- **Indian Institute of Technology, Guwahati** — Guwahati, India
  Bachelor of Technology; Advisor: Amit Sethi — *Jul 2013 - May 2017*

Experience

- **Winterlight Labs** — Toronto, Canada
  Research Engineer, Machine Learning — *Jan 2019 - Aug 2020*
  - **Machine Learning Research Team**: Worked in the machine learning research team on the detection and monitoring of cognitive impairment and mental health conditions from speech.
- **Philips Innovation Campus** — Bengaluru, India
  Research Intern — *May 2017 - July 2017*
  - **Research and Development Division**: Worked in the research division on the application of signal processing and machine learning to healthcare. Projects involved identifying personal information in unstructured medical text for de-identification with neural NER models and personalized sequence prediction of patient health data using LSTMs and sequence modeling.
- **Technische Universitat Dresden** — Dresden, Germany
  Visiting Scholar — *May 2016 – July 2016*
  - **Computer Vision Lab & Computer Graphics Lab**: Project involved visualization of Convolutional Neural Networks (CNNs) at a layer level in 3D by particle rendering. Developed gradient-based importance measure for each filter with respect to output classification in a CNN (using Caffe)
- **Indian Institute of Science** — Bengaluru, India
  Research Intern — *May 2015 – July 2015*
  - **Department of Electrical & Communication Engineering:**: Project involved implementation of various machine learning algorithms in Theano. Analyzed how information is captured in Convolutional Neural Networks using entropy calculations.

Peer-reviewed Publications

- **Balagopalan, A.**, Eyre, B., Rudzicz, F., and Novikova, J., "To BERT or Not To BERT: Comparing Speech and Language-based Approaches for Alzheimer's Disease Detection", 21st Annual Conference of the International Speech Communication Association, INTERSPEECH, 2020
- **Balagopalan, A.**, Novikova, N., McDermott, M. B., Nestor, B., Naumann, T. J., and Ghassemi, M., "Cross-Language Aphasia Detection using Optimal Transport Domain Adaptation", To appear in Machine Learning for Health (ML4H) at NeurIPS 2019, In Machine Learning for Health Workshop (pp. 202-219). PMLR
- Seifert, C., Aamir, A., **Balagopalan, A.**, Jain, D., Sharma, A., Grottel, S., Gumhold, S., "Visualizations of deep neural networks in computer vision: A survey", In Transparent Data Mining for Big and Small Data (pp. 123-144). Springer, Cham., 2017

Peer-reviewed Workshop Papers

- **Balagopalan, A.**, Shkaruta, K., and Novikova, J., "Impact of ASR on Alzheimer's Disease Detection: All Errors are Equal, but Deletions are More Equal than Others", 6th Workshop on Noisy User Generated Text, EMNLP 2020 (Oral Spotlight)
- Novikova, J., **Balagopalan, A.**, Shkaruta, K., Rudzicz, F., "Lexical Features Are More Vulnerable, Syntactic Features Have More Predictive Power", 5th Workshop on Noisy User Generated Text, EMNLP 2019
- **Balagopalan, A.**, Novikova, J., Rudzicz, F., Ghassemi, M., "The Effect of Heterogeneous Data for Alzheimer's Disease Detection from Speech.",Machine Learning for Health (ML4H) at NeurIPS 2018

Clinical Abstracts

- **Balagopalan, A.**, Kaufman, L. J., Novikova, J., Siddiqui, O., Paul, R., Ward, M. and Simpson, W., "Early development of a unified, speech and language composite to assess clinical severity of Frontotemporal Lobar Degeneration (FLTD)", Clinical Trials in Alzheimer's Disease (CTAD) 2019
- Simpson, W., **Balagopalan, A.**, Kaufman, L. J., Yeung, A., and Butler, A., "The Use of a Voice-based Digital Biomarker in Patients With Depression ", International Society for CNS Clinical Trials and Methodology 2019
- **Balagopalan, A.**, Yancheva, M., Novikova, J. and Simpson, W., 2019, "Using acoustic and linguistic markers from spontaneous speech to predict scores on the Montreal Cognitive Assessment (MoCA). ", Memory, 20, p.13., 2019

Projects

- **Inferring task-specific information in text datasets**: Developing an information-theoretic framework to understand how task-specific knowledge, instead of spurious information, is learnable from a text dataset (Oct '20).
- **Cross-language Aphasia Detection**: Proposed neural models for detecting aphasia from a multilingual crowd-sourced speech dataset (Dec '19).
- **Multithreading vs Python GIL: A study**: Project involved studying the effects of the Python Global Interpreter Lock on libraries commonly used for machine learning (NumPy and Scikit-learn) by profiling for concurrency, locks, hotspots of time etc. An adaptive way to set GIL check interval was introduced as a result of profiling. (Jan '18)
- **Using Label Semantics for Image Classification**: Proposed and empirically validated optimization functions for image classification networks to account for relation or similarity between class labels (Dec '17)
- **Predicting the Aesthetic Quality Depicted in an Image Using Deep Learning Approaches**: Use of recurrent attention models for aesthetic rating and training multi-task learning models to identify emotion invoked in a person when viewing an image using the AADB dataset.(Dec '16)

Honors and Awards

- DeepMind PhD Fellowship, University of Toronto 2020-21.
- ISCA Student Grant, University of Toronto 2020.
- MITACS Accelerate Scholarship, University of Toronto from May 2018-Dec 2018.
- Institute Silver Medal, Indian Institute of Technology Guwahati in 2017.
- DAAD Working Internships in Science and Engineering (WISE) Scholarship for a research internship under the guidance of Dr. Carsten Rother, Dr. Stefan Gumhold and Dr. Christin Seifert at TU Dresden
- Kishore Vaigyanik Protsahan Yojana (KVPY) scholarship for higher studies in Basic Sciences, 2013

Invited Talks

- Impact of ASR on Alzheimer's Disease Detection, 6th Workshop on Noisy User Generated Text, EMNLP 2020

Skills Summary

- **Programming Languages**: Python, C++, CUDA, MATLAB, Java
- **Deep Learning Frameworks**: PyTorch, Caffe

Teaching Experience

Teaching Assistant, Department of Computer Science, University of Toronto

- CSC 458: Computer Networks, Fall 2017
- CSC 358: Introduction to Computer Networks, Winter 2018

Extra Curricular

- **DAAD Young Ambassador India 2016-17**: Duties included helping students learn more about research opportunities in Germany
- **Editorial Team Member, InPhase Magazine, 2016**: Was on the editorial team of the departmental magazine, InPhase, in 2016 at IIT Guwahati.
- **Volunteer**: Served as a member in National Service Scheme, 2014.
- **Member**: Electronics Club, IIT Guwahati (2015-16).