Disha Shrivastava

dishu.905@gmail.com | shrivastavadisha.github.io | shrivastavadisha | disha-shrivastava-8398a212

Education

Mila, Université de Montréal, Canada — Sept 2018 - Present
PhD in Machine Learning — *GPA: 4.3/ 4.0*

Indian Institute of Technology Delhi, India — July 2014 - 2016
MTech in Computer Technology — *CGPA: 9.44/ 10.0*

Birla Institute of Technology Mesra, India — May 2008 - 2012
B.E. in Electronics & Communication Engg. — *CGPA: 8.64/ 10.0 (abs)*

Work Experience

Google Brain — Aug 2019 - PRESENT
Student Researcher — *Montreal, Canada*
Developing machine learning models for adaptation in program understanding and generation.

IBM Research — Aug 2016 - Aug 2018
Research Software Engineer — *Bangalore, India*
KGs for domain-specific data, Reasoning for complex QA, Metrics for computational creativity and topical coherence.

Idea Cellular Limited — July 2012 - Nov 2013
Assistant Manager — *Hyderabad, India*
Configuration and O&M of nodes responsible for real-time charging and billing of prepaid mobile subscribers.

Publications and Patents

[1] Disha Shrivastava, Hugo Larochelle, Daniel Tarlow. **Learning to Combine Per-Example Programs for Neural Program Synthesis** (*NeurIPS 2021*)

[2] Edoardo Maria Ponti*, Rahul Aralikatte*, Disha Shrivastava, Siva Reddy, Anders Søgaard. **Minimax and Neyman-Pearson Meta-Learning for Outlier Languages** (*Findings of ACL 2021*)

[3] Disha Shrivastava, Hugo Larochelle, Daniel Tarlow. **On-the-Fly Adaptation of Source Code Models** (*NeurIPS Workshop on Computer-Assisted Programming 2020*)

[4] Disha Shrivastava*, Eeshan Gunesh Dhekane*, Riashat Islam. **Transfer Learning by Modeling a Distribution over Policies** (*ICML Workshop on Multi-Task and Lifelong Reinforcement Learning 2019*)

[5] Disha Shrivastava, Saneem Ahmed CG, Anirban Laha, Karthik Sankaranarayanan. **A Machine Learning Approach for Evaluating Creative Artifacts** (*SIGKDD Workshop on Machine Learning for Creativity 2017*)

[6] Disha Shrivastava, Santanu Chaudhury, Dr. Jayadeva. **A Data and Model-Parallel, Distributed and Scalable Framework for Training of Deep Networks in Apache Spark** (*arXiv 2017*)

[7] Disha Shrivastava, Sreyash Kenkre, Santosh Penubothula. **Hypernyms through Intra-Article Organization in Wikipedia** (*arXiv 2018*)

[8] Disha Shrivastava, Abhijit Mishra, Karthik Sankaranarayanan. **Modeling Topical Coherence in Discourse without Supervision** (*arXiv 2018*)

[9] Shavak Agrawal, Anush Sankaran, Anirban Laha, Saneem Ahmed CG, Disha Shrivastava, Karthik Sankaranarayanan. **What is Deemed Computationally Creative?** (*IBM Journal of Research and Development 2019*)

[10] Pankaj S Dayama, Disha Shrivastava. **System and Method to Implement a Cognitive Quit Smoking Assistant** (*US Patent App. 15/811, 964*)

[11] Anirban Laha, Vijay Ekambaram, Parag Jain, Disha Shrivastava. **Displaying Dynamic Content on Multiple Devices** (*US Patent 10,664,217*)

[12] Sreyash Kenkre, Santosh R.K. Penubothula, Disha Shrivastava, Harish Guruprasad Ramaswamy, Vinayaka Pandit. **Automated Constraint Extraction and Testing** (*US Patent 10,902,200*)

[13] Anush Sankaran, Pranay Lohia, Priyanka Agarwal, Disha Shrivastava, Anirban Laha, Parag Jain. **Cognitive Assistant for Co-Generating Creative Content** (*US Patent App. 16/169,001*)

Projects

Learning to Combine Per-Example Solutions using Neural Program Synthesis
Nov 2020 - May 2021

Advisors: Hugo Larochelle and Danny Tarlow — *Mila/ Google Brain, Canada*

- Formulated the standard program synthesis pipeline as a two stage process: (a) finding programs that satisfy a single example (PE solutions) (b) aggregating the PE solutions such that it leads to a global solution.
- Proposed a neural network based multi-head attention architecture called Cross Aggregator that uses the step-wise partial execution information of the programs to learn to combine the PE solutions.
- Demonstrated significant improvements in performance over two different experimental settings and different program lengths.

On-the-Fly Adaptation of Source Code Models
Dec 2018 - Oct 2020

Advisors: Hugo Larochelle and Danny Tarlow — *Mila/ Google Brain, Canada*

- Formulating the task of code auto-completion in an IDE for a new file, in terms of inner-loop adaptation using targeted information (support tokens) from the local context.
- Demonstrated significant benefits in cross-entropy and MRR over baselines including dynamic evaluation, with huge improvements in case of identifiers and literals.

Minimax and Neyman-Pearson Meta-Learning for Outlier Languages
Oct 2020 - Feb 2021

Collaborators: Edoardo Maria Ponti*, Rahul Aralikatte*, Siva Reddy and Anders Søgaard — *Mila, Canada*

- Proposed two variants of MAML based on alternative criteria: (a) Minimax MAML: reduces the maximum risk across languages, (b) Neyman–Pearson MAML: constrains the risk in each language to a maximum threshold. Both criteria constitute fully differentiable two-player games that are more suited to outlier languages.
- Demonstrated gains in average and minimum performance across low-resource languages in zero and few-shot settings for POS tagging and QA, when compared to joint multisource transfer and vanilla MAML.

Transfer Learning by Modeling a Distribution over Policies
Feb 2019 - June 2019

Collaborators: Eeshan Gunesh Dhekane and Riashat Islam — *Mila, Canada*

Built on the idea of modeling a distribution over policies in a Bayesian deep reinforcement learning setup to propose a transfer strategy which leads to faster exploration in the target environment by maximizing the entropy of a distribution of policies.

Knowledge Graph Construction and Reasoning for Domain-Specific Data
Aug 2016 - Aug 2018

Collaborators: Vinayaka Pandit, Sreyash Kenkre and Indrajit Bhattacharya — *IBM Research, India*

- Actively involved in developing and analyzing an end-to-end ***unsupervised framework for open-domain Knowledge Graphs (KG) construction for domain-specific datasets***. The framework takes text corpus of the specific domain along with some meta-data from Wikipedia as input and gives domain and document KGs, sentence-wise annotated concepts, relations and triples along with their domain-wise importance scores and a set of connected Wikipedia Categories as outputs.
- Worked to further improve the quality of the KG by doing entity and relation canonicalization and linking. Developed a novel unsupervised and computationally light technique for ***Hypernym Detection and Directionality using the structure of documents***.
- Worked towards increasing the usability by exposing this framework to other teams within IBM in form of micro-services API. Our specific target domains included ***Financial and Service Compliance Documents and IT System Logs***.
- Developed and implemented a framework which facilitates reasoning over the KG formed to retrieve a ranked list of paragraphs for the task of ***complex Question-Answering***.
- Worked towards ***generating programs for arithmetic problems*** which can serve as explanations (hence more interpretable) and which when executed can produce the correct answer similar in spirit to Neural Program Interpreters.

Machine Learning Approaches for Evaluating Creative Artifacts
Apr 2017 - Sept 2017

Collaborators: Karthik Sankaranarayanan, Saneem Ahmed CG, Anirban Laha — *IBM Research, India*

- Contributed towards postulating the ***dimensions and factors that distinguish computational creativity*** and intelligence, such as novelty, value, surprise, influence, coherence, correctness, and comprehensibility. The application domains are grouped into time-dependent and time-independent ones and framework is defined to describe these dimensions in each application.
- Incorporating important measures for creativity (e.g. novelty, influence, unexpectedness, value, etc.), proposing a ***regression-based learning framework for evaluating these metrics*** and analyzing the results in the domain of movies leading to improvement in prediction of both critic and audience movie ratings.

Modelling Topical Coherence in Discourse without Supervision
Nov 2017 – March 2018

Collaborators: Abhijit Mishra and Karthik Sankaranarayanan — *IBM Research, India*

Developed an interpretable, unsupervised metric to come up with a coherence score for an input text paragraph. It relies on extracting topics from all sentences followed by: (a) measuring the degree of uncertainty of the topics with respect to the paragraph, and (b) measuring the relatedness between these topics using the graph structure. Experiments on a public essay dataset and synthetic dataset show positive correlation with the ground-truth as well as significant agreement with human judges.

Large Scale Distributed Deep Learning
July 2015 - July 2016

Advisors: Prof. Santanu Chaudhury and Prof. Jayadeva — *IIT Delhi, India*

- Developed a distributed and scalable framework for efficient training and inference of a generic deep neural network architecture (Fully-Connected Feedforward Networks, CNN, Autoencoders, RNN and LSTM) implementing both ***Data Parallelism and Model Parallelism over a cluster of cheap commodity hardware (CPUs) using Apache Spark***.
- Proposal of a new algorithm for training of deep networks for the case when the network is partitioned across the machines (Model Parallelism); along with detailed cost analysis and mathematical and experimental proof of convergence of the algorithm.
- Applied the developed framework for ***noise resilient image super-resolution*** beating state of art techniques in terms of both PSNR and SSIM with significant reduction in training time and improved scalability.
- Achieved ***11X speedup for 5M samples*** and ***5.6X speedup with 4 billion model parameters*** for CNN; and ***7.2X speedup for 0.1M samples*** for Fully-Connected Nets [baseline: one machine] over a cluster of just ***5 CPU's*** - no GPU.

Classification Algorithms for EEG based Brain Computer Interfaces JAN 2015 - MAY 2015

ADVISOR: PROF. JAYADEVA *IIT Delhi, India*

- Implemented and analyzed different feature selection and classification algorithms on ***BCI Motor Imagery*** data.
- Achieved accuracy (kappa = 0.4) ***equivalent to the winner*** of the BCI Competition IV (dataset 2b).

Technical Skills

Programming Languages C, C++, Java, Python, Scala, Bash, VHDL

Softwares and Packages TensorFlow, PyTorch, Keras, Caffe, Apache Spark, Hadoop, MySQL, MATLAB, OpenCV

Positions of Responsibility

Workshops Organized
- Advances in Programming Languages and Neurosymbolic Systems (AIPLANS) to be held at NeurIPS 2021
- 1st International Workshop on Machine Learning for Creativity held at ACM SIGKDD 2017 at Halifaux, Canada

Reviewer ICLR 2022, NeurIPS 2020-21, ICML 2020, ACL 2020, AAAI 2020-21, MAIS 2019, GHCI 2017 (AI and ML Track)

Teaching Assistant Pattern Recognition (July-Nov 2015), Software Lab (Jan-May 2016) at IIT Delhi

Others
- Served as a mentor for AI4Good Lab 2021.
- Volunteer at WiML workshop at NeurIPS 2019.
- Part of Diversity and Inclusion initiatives at IBM Research, India (July 2017 - July 2018).
- Student Placement Coordinator at Training and Placement Cell, BIT Mesra (July 2011 - Apr 2012).
- Organizer and convener of the ML4Code reading group at Mila (Jan 2021 - present).

Selected Courses

Mila Probabilistic Graphical Models, Reinforcement Learning

IIT Delhi Deep Learning for Vision, Neural Networks, Computer Vision, Operating Systems, Database Management Systems, Computer Architecture, System Software, Software Lab, Coding Theory, Computer Networks, Computer Networks Lab

BIT Mesra Digital Signal Processing, Data Structure in C++, Unix & C Programming, Random & Stochastic Process, Digital Image Processing, Digital Electronics, Microprocessor & Interfacing, Linear Control Theory