

# Disha Shrivastava

✉ dishu.905@gmail.com | 🌐 shrivastavadisha.github.io | 📄 shrivastavadisha | 🐦 @DishaShrivasta9

## Education

### Mila, Université de Montréal, Canada,

PHD IN MACHINE LEARNING, SUPERVISORS: HUGO LAROCHELLE AND DANNY TARLOW

SEPT 2018 - PRESENT

GPA: 4.3/ 4.0

### Indian Institute of Technology Delhi, India

MTECH IN COMPUTER TECHNOLOGY

JULY 2014 - 2016

CGPA: 9.44/ 10.0

### Birla Institute of Technology Mesra, India

B.E. IN ELECTRONICS & COMMUNICATION ENGG.

MAY 2008 - 2012

CGPA: 8.64/ 10.0 (abs)

## Experience

### Google Brain

STUDENT RESEARCHER, SUPERVISORS: HUGO LAROCHELLE AND DANNY TARLOW

Incorporating repository-level context in LLMs of source code.

NOV 2022 - PRESENT

Montreal, Canada

### DeepMind

RESEARCH SCIENTIST INTERN, COLLABORATORS: YUJIA LI AND DAVID CHOI

AlphaCode

JUN 2022 - OCT 2022

London, UK

### Google Brain

STUDENT RESEARCHER, COLLABORATORS: HUGO LAROCHELLE, DANNY TARLOW AND CHARLES SUTTON

Machine learning models for adaptation in program understanding and generation.

AUG 2019 - JUN 2022

Montreal, Canada

### IBM Research

RESEARCH SOFTWARE ENGINEER

KGs for domain-specific data, reasoning for complex QA, metrics for computational creativity and topical coherence.

AUG 2016 - AUG 2018

Bangalore, India

### Idea Cellular Limited

ASSISTANT MANAGER

Configuration and O&M of nodes responsible for real-time charging and billing of prepaid mobile subscribers.

JULY 2012 - NOV 2013

Hyderabad, India

### Indian Institute of Technology Kharagpur

RESEARCH INTERN, SUPERVISOR: JAYANTA MUKHOPADHYAY

Advanced restoration and enhancement of color images

MAY 2011 - JUN 2011

Kharagpur, India

### Indian Institute of Technology Kanpur

RESEARCH INTERN, SUPERVISOR: PRADEEP KUMAR K

Effects of chromatic dispersion in fiber optic systems.

MAY 2010 - JULY 2010

Kanpur, India

## Publications and Patents

Disha Shrivastava, Hugo Larochelle, Daniel Tarlow. **Repository-Level Prompt Generation for Large Language Models of Code** (ICML Workshop on Knowledge Retrieval and Language Models 2022)

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

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

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

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)

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)

[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*)

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

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

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*)

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

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

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

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*)

## Positions of Responsibility

---

### Organization

- Deep learning for Code (DL4C) workshop held at ICLR 2022
- Advances in Programming Languages and Neurosymbolic Systems (AIPLANS) workshop held at NeurIPS 2021
- Machine Learning for Creativity workshop held at ACM SIGKDD 2017
- ML4Code reading group at Mila

### Mentoring

- Project: *Inducing reasoning in language models of smaller scale*  
Course: Natural Language Understanding with Deep Learning/ Computational Semantics, Fall 2022, McGill  
Student: Pulkit Madan
- Project: *ML tool for diagnosis of a patient's underlying cause of dysarthria by classifying audio input as being indicative of Parkinson's disease, ALS or Cerebral Palsy*  
Initiative: AI4Good Lab, 2021  
Students: Chloe Pappas, Ritu Ataliya, Nadia Enhaili, Hala Hassan, Jiayue Yang and Kamun Karl Itaj.
- Project: *Towards systematicity in seq2seq models*  
Course: Natural Language Understanding with Deep Learning, Winter 2020, McGill  
Students: Dora Jambor and Emily Goodwin

### Teaching Assistant

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

### Reviewer

ICML 2022, ICLR 2022, NeurIPS 2020-22, ICML 2020, ACL 2020, AAAI 2020-21, MAIS 2019, GHCI 2017 (AI and ML Track)

### Others

- 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).

## Invited Talks

---

### Repository-Level Prompt Generation for Large Language Models of Code

- Compiler Reading Group, Google (online)
- Program Synthesis Reading Group, DeepMind (London)
- Tea-Talk, Google Brain (Montreal)
- ML4Code Reading Group, Mila (Montreal)

### Learning to Combine Per-Example Solutions for Neural Program Synthesis

- Neurosymbolic Group, MIT (online)
- Tea-Talk, Google Brain (Montreal)
- ML4Code Reading Group, Mila (Montreal)

### On-the-fly Adaptation of Source Code Models

- Ada Support (Toronto)
- Tea-Talk, Google Brain (Montreal)

## Awards and Achievements

---

- Summer Undergraduate Research Award: Among 40 students across India to get SURGE 2010 Fellowship by IIT Kanpur.
- Secured an All India Rank of 189 out of 216367 candidates in GATE 2014.
- Secured the 1st position in Bachelor of Engineering across all branches ( 500 students) in 3rd, 6th, 7th and 8th semesters

- Division Topper in Class XII CBSE.
- Secured the 2nd position in Masters in Computer Technology.
- Recipient of ICML 2019 Diversity & Inclusion Travel Grant.

## Research

### LLM-based Interactive Framework for the Task of Software Documentation

OCT 2022 - PRESENT

COLLABORATORS: JIN L.C. GUO AND AVINASH BHAT

*Mila, Canada*

- Conducting a user-study to analyze and encode the interaction of tutorial-writers with LLMs as well as identify the challenges that they face.
- Developing a LLM that can serve as an interactive tool to overlap potential challenges identified above by incorporating user feedback.
- Validating the effectiveness of the developed framework by conducting user-study with the same set of participants.

### Repository-Level Prompt Generation for Large Language Models of Code

NOV 2021 - MAY 2022

ADVISORS: HUGO LAROCHELLE AND DANNY TARLOW

*Mila/ Google Brain, Canada*

- Proposed a framework that learns to generate example-conditioned prompts without requiring access to the weights of the LLM. The framework enables generating prompts using domain-knowledge, in the form of structure as well as the context from all files in the repository.
- Demonstrated significant improvement over Codex and other baselines on the task of single-line code-autocompletion.

### 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: SANEEM AHMED CG, ANIRBAN LAHA AND KARTHIK SANKARANARAYANAN

*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: SANTANU CHAUDHURY AND 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: 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).

## Miscellaneous

---

- My short story (Journey for Justice) was published as part of a book entitled "Blank Space" which was released at the New Delhi World Book Fair 2015 held at Pragati Maidan, New Delhi, India.
- I developed a legged robot using AVR Microcontroller as part of Arthrobotix: a workshop conducted by Technophilia at BIT Mesra, India.
- I developed an Aerobot as part of a workshop on Aerial Robotics conducted by MECHAHAWKS at BIT Mesra.
- I performed Kathak (indian classical dance) for SPIC MACAY under the mentorship of Pandit Rajendra Kumar Gangani and his team at Dogra Hall, IIT Delhi, India.
- I was part of the organizing committee of Unnayan 2009: Inter College Technical Fest organized by IEEE Students Chapter, BIT Mesra, India.
- I was an active member of News and Publication Society, BIT Mesra, India.