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*
Supervisors: Hugo Larochelle and Danny Tarlow

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

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

Work Experience

Google Brain — Aug 2019 - PRESENT
Student Researcher — *Montreal, Canada*
Working on applications of meta-learning for adaptation of neural models based on local context

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

[2] 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*)

[3] 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*)

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

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

[6] 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*)

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

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

[9] Sreyash Kenkre, Santosh R.K. Penubothula, Disha Shrivastava, Harish Guruprasad Ramaswamy, Vinayaka Pandit. **Automated Constraint Extraction and Testing** (*US Patent App. 16/186,924*)

[10] 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

On-the-Fly Adaptation of Source Code Models — Dec 2018 - Jan 2020
Advisors: Hugo Larochelle and Danny Tarlow — *Mila/ Google Brain, Canada*
Worked on formulating the task of adaptation to local context (file) in source code in terms of inner-loop adaptation using support tokens. We also develop a method for selecting targeted information from the file which is then evaluated in a setting meant to mimic the code auto-completion workflow in an IDE. Experiments show significant benefits in performance over baselines including dynamic evaluation with significant improvements in case of identifiers.

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).

Age and Gender Classification using CNN by Segmenting parts of Face Image

Aug 2015 - Nov 2015

Advisor: Dr. Raghavendra Singh — *IIT Delhi, India*

- Classification of Age and Gender by ***segmenting the eyes, mouth and nose from full face images*** and training a CNN model followed by analysis of results in terms of variation of model parameters and bounding box projection visualization at different layers of the model.
- Achieved an accuracy 3% less than the state-of-art results for age classification for full face images with just the eye-cropped images for ***Adience-OUI*** dataset

Spectrum Sensing in Cognitive Radio Environments

July 2011 - Apr 2012

Advisor: Dr. Sanjay Kumar — *BIT Mesra, India*

Worked towards modeling and carrying out simulations to develop and analyze flexible spectrum usage techniques to increase the throughput for single and multiple secondary user scenarios in cognitive radio environments by incorporating effective algorithms for cooperative spectrum sensing which are independent of distance and time factors.

Short-Term Projects

July 2014 - May 2015

Advisors: Sumantra Dutta Roy and Sumeet Agarwal — *IIT Delhi, India*

- *Computer Vision related projects:* (i) Feature Preserving Morphing of one face into another, (ii) Mosaicing of different images of a scene to form a panaroma, (iii) Motion Segmentation and Object Removal in Videos.
- *System Software related projects*: (i) FTP Server with Multiple Clients using Socket Programming in C, (ii)Sticky notes App in Python: Used Tkinter for GUI, (iii) Shell implementation: Designed a shell (like Bash) in C for executing linux commands via system calls.
- *Computer Architecture related project*: Sorting algorithms on parallel architectures: Merge sort and Quick sort on hypercube and linear ring.
- *Compiler Design related project*: Parser (using lex and yacc) for a grammar with restricted subset of English language.

Scholastic Achievements

- Second rank holder in the class while pursuing Masters in Computer Technology at IIT Delhi
- 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 in 3rd, 6th, 7th and 8th semesters.
- Division Topper in Class XII CBSE
- Recipient of ICML 2019 Diversity & Inclusion Travel Grant

Technical Skills

Programming Languages C, C++, Java, Python, Scala, Bash, VHDL
Softwares and Packages TensorFlow, PyTorch, Keras, Caffe, Apache Spark, Hadoop, MySQL, MATLAB, OpenCV

Internships

Advanced Restoration and Enhancement of Color Images May 2011 - Jun 2011
Advisor: Jayanta Mukhopadhyay, Department of CSE, IIT Kharagpur *IIT Kharagpur, India*
Developed a GUI with multiple tools in Matlab for Advanced Image Restoration and Improving Image Quality (PhotoEditing)

Effects of Chromatic Dispersion in fiber optic systems May 2010 – July 2010
Advisor: Dr. Pradeep Kumar K, Dept. of Electrical Engg., IIT Kanpur *IIT Kanpur, India*
Developed models and analyzed the BER performance of single and multichannel 10 Gb/s RZ-DPSK systems using Simulink

Positions of Responsibility

Workshops Organized 1st International Workshop on Machine Learning for Creativity held at ACM SIGKDD 2017 at Halifaux, Canada
Reviewer NeurIPS 2020, 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
- Volunteer at WiML workshop at NeurIPS 2019.
- Student Placement Coordinator at Training and Placement Cell, BIT Mesra (July 2011 - Apr 2012)
- Part of Diversity and Inclusion initiatives at IBM Research, India (July 2017 - July 2018)

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

Extracurricular Activities

- Part of the organizing committee of Unnayan 2009: Inter College Technical Fest organized by IEEE Students Chapter, BIT Mesra, India.
- Active member of News and Publication Society, BIT Mesra, India.
- Developed a legged robot using AVR Microcontroller as part of Arthrobotix: a workshop conducted by Technophilia at BIT Mesra, India.
- Developed an Aerobot as part of a workshop on Aerial Robotics conducted by MECHAHAWKS at BIT Mesra.
- Kathak (indian classical dance) performance for SPIC MACAY under the mentorship of Pandit Rajendra Kumar Gangani and his team at Dogra Hall, IIT Delhi, India.
- 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.