

Prateek Varshney

· Google Research India · Indian Institute of Technology Kanpur

☎ (+91) 9868588504 | ✉ p.varshney1729@gmail.com | 🏠 pvarshney1729.github.io | 📱 pvarshney1729 | 📄 pvarshney1729

Education

Indian Institute of Technology Kanpur

Kanpur, India

BACHELOR OF TECHNOLOGY, MAJOR IN COMPUTER SCIENCE AND ENGINEERING

2017 - 2021

MINOR IN COGNITIVE SCIENCE | CUM. GPA: 9.5/10.0

Hillwoods Academy

Delhi, India

ALL INDIA SENIOR SECONDARY CERTIFICATE EXAM (AISSCE) | PERCENTAGE: 95.8%

2016

Publications

CS-NET at SemEval-2020 Task 4: Siamese BERT for ComVE

[Paper] [Code]

SOUMYA RANJAN DASH, SANDEEP ROURAY, PRATEEK VARSHNEY, ASHUTOSH MODI

Dec 2020

- Proposed a system for Task 4: "Commonsense Validation and Explanation" of SemEval 2020 which involved differentiating between natural language statements that confirm to common sense and those that do not
- Developed a Siamese model based on transformer neural network architecture, which was able to select the against common sense statement and identify the most crucial reason why a statement does not make sense
- Paper accepted at International Workshop on Semantic Evaluation 2020

Honors & Awards

2021 **EPFL E3 Program**, Among top 100 students in India

Switzerland

2020 **S.N. Bose Scholarship**, Among top 25 students in India

USA

2018 **Academic Excellence Award**, For exceptional academic performance in 2017-19

IIT Kanpur

2017 **Joint Entrance Exam (JEE) Advanced**, Top 1.13% among 0.17 million candidates

India

2017 **Joint Entrance Exam (JEE) Mains**, Top 0.04% among 1.2 million candidates

India

2015 **KVPY Scholarship**, Indian Institute of Science and Government of India

Bangalore

2014 **All India Rank 54**, CBSE Group Mathematical Olympiad

India

Work Experience

Google Research India

Bangalore, India

RESEARCH ASSOCIATE UNDER DR. PRATEEK JAIN

Jun 2021 - Present

- Developing differential privacy framework to analyse and improve the estimator risk vs privacy budget tradeoffs of stochastic optimisation routines in fundamental machine learning problems such as Linear Regression Settings
- Proposed DP-SGD and Momentum based routines have improved risk guarantee. than current state of the art results.
- Working on Sample Efficient Meta-Learning using Alternate Minimization of Linear Regression Tasks with sparse vectors.

École polytechnique fédérale de Lausanne

Lausanne, Switzerland

E3 INTERN UNDER PROF. VOLKAN CEVHER AND DR. GRIGORIOS CHRYSOS

Jul 2021 - Present

- Working on combining Unsupervised Domain Adaptation (UDA) and Class Incremental (CI) learning methods to handle both Domain Shift Problem and Difference of Classes in the Training and Target Space and improve the generalizability
- Able to achieve high accuracies with MNIST->MNIST-M/SVHN Settings using a novel combination of a trainable embedding module, conditional GAN, class-specific representative anchor points and negative space concept

California Institute of Technology

Pasadena, California

RESEARCH SCHOLAR UNDER PROF. ANIMASHREE ANANDKUMAR AND DR. FLORIAN SCHAFFER

Aug 2020 - Jan 2021

- Worked on the intersection of Game Theory/Mechanism Design and Deep Learning and developed zero-sum game strategies between the Discriminator and an Adversary to solve inter and intra-class imbalance
- Obtained robust classifier models for problems in Computer Vision and Protein Function Prediction

Microsoft India (R&D), Bangalore

Bangalore, India

SOFTWARE ENGINEERING INTERN

May 2020 - Jul 2020

- Implemented an ADF pipeline for Demand Forecasting, which loaded inventory snapshots & deltas from ADLS Gen 2 storage, performed suitable transformations and various statistical analytics algorithms in Databricks
- Linked the ADLS Gen 2 storage to Azure Data Share and enabled scheduled cross subscription sharing of data snapshots Created an interactive dashboard using PowerBI for visualisation of the predictions and insights obtained
- Received a Pre-Placement Offer for my work during course of the internship

National University of Singapore

Kent Ridge, Singapore

VISITING RESEARCH SCHOLAR UNDER PROF. DJORDJE JEVDJIC

May 2019 - Jul 2019

- Designed an open-source DNA-based archival storage tool. Implemented a distributed subquadratic clustering algorithm which converged efficiently on real/synthetic DNA datasets and was robust to outliers and high levels of noise
- Simulated storing and reconstructing error-sensitive compressed image files from noisy amplified DNA strands

Selected Projects

Speaker Diarization

[Paper] [Code]

PROJECT UNDER PROF. VIPUL ARORA

Feb 2021 - Apr 2021

- Developed a Speaker Diarization Model with Transfer Learning Flavour
- Created an end to end pipeline by leveraging Transfer Learning on top of an LSTM-based text-independent speech embedding model, which is passed to a parametric clustering algorithm to obtain a speaker diarization system
- Experimented with various VADs, Pre-processing Algorithms, Embedding Modules and Clustering Algorithms, and empirically verified the potential of Transfer Learning in improving the training time at very negligible accuracy loss

Control Variates for Stochastic Gradient Hamiltonian Monte Carlo

[Paper] [Code]

PROJECT UNDER PROF. DOOTIKA VATS

Jan 2021 - Apr 2021

- Explored the avenues of variance reduction methods such as Control Variates and their applications to Stochastic Gradient based Langevin Dynamics (SGLD), MCMC (SGMCMC) and Hamiltonian Monte Carlo (SGHMC) techniques
- Reproduced the results of two papers: "Variance Reduction for Stochastic Gradient Optimisation" and "Control Variates for Stochastic Gradient MCMC" and explored their theoretical aspects
- Extended Control Variates to different settings such as Metropolis-adjusted Langevin algorithm (MALA)

Conditional Random Field model for ATAC-seq data

[Slides][Code]

CS690A COURSE PROJECT UNDER PROF. HAMIM ZAFAR

Sep 2020 - Dec 2020

- Implemented a Hidden Markov Model (HMM) for ATAC-seq data to identify open chromatin regions
- Improved upon the baseline HMM performance with a Conditional Random Field (CRF) and NLP inspired models

Comparison of SGD Variants for Stochastic Optimization

[Paper] [Code]

EE609A TERM PROJECT UNDER PROF KETAN RAJAWAT

Mar 2020 - May 2020

- Reproduced and extended the results of "On the Insufficiency of Existing Momentum Schemes for Stochastic Optimization" and "Accelerating Stochastic Gradient Descent For Least Squares Regression" by Kidambi et al.
- Showed experimentally that there exist simple stochastic problem instances where momentum based methods are sub-optimal and enjoy practical gains over SGD in deep learning applications due to minibatching
- Established that ASGD and Adam can converge faster than all other methods irrespective of batch sizes

Smart Tutor

[Code]

PROJECT UNDER PROF. VIPUL ARORA

Dec 2018 - Feb 2019

- Worked on a review scheduling algorithm, using model-free reinforcement learning to learn a teaching policy
- Studied and contrasted three student learning environments: Ebbinghaus' Curve, Half-life Regression, and Generalised Power Law; two teaching performance metrics and four baseline policies

Skills

Programming	C/C++, Python, R, Haskell, GNU Octave, Node.js, HTML, CSS, MySQL, MongoDB
Deep Learning	Tensorflow, Pytorch, Scikit-Learn, Pillow, Keras, Numpy, Pandas, Matplotlib, Scipy, CNTK, OpenCV
Utilities	Linux shell utilities, Git, Bash, GDB, \LaTeX

Relevant Coursework

Introduction to Machine Learning	Introduction to Programming	Probability for Computer Science
Special Topics in NLP	Data Structures and Algorithms	Real Analysis & Multivariate Calculus
Convex Optimization in SP-COM	Algorithms II	Discrete Mathematics
Speech Signal Processing ⊙	Theory of Computation	Linear Algebra and ODE
Machine Learning for SP ⊙	Modern Cryptology	Stat. Simulation & Data Analysis A★
Advanced Topics in ML	Operating System	Markov Chain Monte Carlo A★
Computational Genomics	Linear Algebra and ODE	Computational Physics
Probabilistic Machine Learning	Compilers	Introduction to Bayesian Analysis ⊙

⊙ ≡ Audit, A★ ≡ Exceptional Performance (Top 1%), **bold** ≡ Graduate Level Course

Mentorship Roles and Extra Curriculars

Teaching Assistant INTRODUCTION TO COMPUTING, PROF. DEBADUTTA MISHRA

Aug 2020 - Nov 2020,

AND PROF. BISWABANDAN PANDA

Jan 2021 - May 2021

- Developed programming assignments, designed evaluation metrics and assisted in checking of papers and quiz copies

Academic Mentor COUNSELLING SERVICE, IIT KANPUR

May 2018 - Jul 2019

- Mentored a Batch of 900+ freshmen in Introduction to Computing course via Institute Level Remedial Classes
- Conducted Doubt Clearing Sessions and provided intensive Individualized Mentoring to special needs students from diverse socio-economic backgrounds including limited English/Hindi-speaking students

Core Group Member VOX POPULI

Oct 2018 - Apr 2021

- Team member of Vox Populi, the campus journalism society of IIT Kanpur