Rohan Shah

CSE Undergraduate, IIT Bombay

rohanshah1306@gmail.com | github.com/rohanshah13

Education

Indian Institute of Technology Bombay *2018-2022*
Bachelors of Technology with Honors in Computer Science and Engineering *CPI: 9.79 / 10.00*

Publications

- **PAC Mode Estimation using PPR Martingale Sequences** [Paper]
  Shubham Jain; **Rohan Shah***; Sanit Gupta*; Denil Mehta*; Inderjeet Nair*; Jian Vora*; Sushil Khyalia; Sourav Das; Vinay Ribeiro; Shivaram Kalyanakrishnan;
  *(Under review at AISTATS 2022, * indicates equal contribution)*

Research Experience

PAC Mode Estimation using PPR Confidence Sequences *Jan 2021 - Oct 2021*
Prof. Shivaram Kalyanakrishnan *RnD Project, IIT Bombay*

- Used the prior-posterior ratio (PPR) Martingale Test as a computationally efficient stopping rule for mode estimation, observed that it is more sample-efficient than all baselines
- Proved that the sample complexity of PPR is optimal in the asymptotic regime, i.e. as the mistake probability goes to 0, the first such guarantee for mode estimation
- Established that the one-vs-one application of PPR for multi-class mode estimation is provably more sample efficient than the one-vs-rest and multi-dimensional Dirichlet variants

Winner Forecasting in Indian Elections *Jan 2021 - Sep 2021*
Prof. Shivaram Kalyanakrishnan *RnD Project, IIT Bombay*

- Proposed the novel bandit inspired DCB (Difference in Confidence Bounds) algorithm to decide which constituency to poll in order to efficiently predict a two level election (held in India, UK)
- Formulated an Integer Programming problem to determine the optimal choice of constituencies in the presence of perfect information, allowing the computation of a ”regret” term
- Implemented efficient large scale simulations on real world and synthetic election datasets, observed that the sample complexity of DCB is within a factor of 1.5 of the optimal value

Zero-Shot Cross Lingual Transfer using Adapters *Jul 2021 - present*
Prof. Preethi Jyothi *Bachelor's Thesis Project, IIT Bombay*

- Utilized several source language adapters to fit BERT to unseen target languages at test time
- Experimentally determined that language similarity (in terms of syntax, geographic proximity etc) is not a good criterion to choose source language adapters for cross-lingual transfer
- Discovered that transfer languages can be chosen based on performance on labelled english data, regardless of target languages

Black-Box Error Detection in QnA systems *Jan 2021 - Sep 2021*
Prof. Zhendong Su *Research Internship, ETH Zurich*

- Investigated the nature and frequency of mistakes in results given by Google and Bing, and their variation with respect to time, browser and operating system using web crawlers
- Proposed differential and metamorphic testing approaches (back-translation, verb perturbations) that eliminate the need for high-quality labelled data to identify bugs
- Conducted experiments that show the proposed methods find bugs with 2x precision as compared to using labelled QnA datasets, with 20+ reported issues fixed by developers
- Studied Google and Bing voice assistants in depth, identifying bugs and unexpected behaviour

Internships

Automated Feedback Loop for Candidate Recommendation System *May 2021 - Jul 2021*
Global Talent Acquisition Team *Microsoft India Development Center*

- Received a full time offer based on performance in the internship
- Proposed and implemented an automated feedback loop for an AI based resume to job description matching system, using techniques from Information Retrieval
- Leveraged Azure cloud storage and machine learning services for an efficient implementation
- Implemented an automated evaluation plan in C# and presented several useful findings

Plumbing Simulation with Intelligent Feedback System *May 2020 - Jul 2020*
Prof. Chandan Dasgupta *Educational Technology, IIT Bombay*

- Developed a multiplayer simulation of a plumbing system for educational purposes using Django Channels powered web sockets and a React frontend
- Enabled a real time chat feature and automated generation of activity reports using PostgreSQL
- Publicly released in schools to help students understand engineering problems and learn to make optimal design choices

Academic Achievements

- Awarded the **Institute Academic Prize** for exceptional academic performance *2019*
- Secured All India Rank **77** in **IIT-JEE** Advanced 2018 amongst **230,000** candidates *2018*
- Secured All India Rank **187** in **JEE** Mains 2018 amongst **1 million+** candidates *2018*
- Among the top **25(0.1%)** included in Training Camp for **International Astronomy Olympiad** *2018*
- Among the **top 12 (1%)** granted **Branch Change** to **Computer Science and Engineering** *2019*
- Awarded AP grades for outstanding (**top 1%**) performance in Digitial Logic Design and Introduction to Electrical Systems *2020*

Technical Projects

Graph Database based Recommendation System *Ongoing*
Prof. Umesh Bellur | Database Management *Course Project*

- Generated real time recommendations using the GraphAware framework, integrated with Neo4j
- Obtained node embedding vectors for each user and movie using the BINE Graph Neural Network
- Tested collaborative and content based recommendation techniques on IMDb dataset of 100k users

Adversarial Attacks on Speech Recognition Systems *Spring 2021*
Prof. Preethi Jyothi | Automatic Speech Recognition *Course Project*

- Studied targeted, imperceptible, white-box and black-box adversarial attacks on speech systems
- Investigated pscho-acoustic hiding (phenomenon in which certain frequencies become inaudible to humans) for better imperceptibility and implemented attacks in Indian languages

Medical Image Segmentation using Recurrent Residual U-Net *Spring 2020*
Prof. Suyash Awate | Medical Image Computing *Course Project*

- Implemented segmentation of skin cancer and retinal blood vessel images in TensorFlow, performed an ablation study on the effect of recurrent and residual connections
- Achieved competitve results on the ISIC and STARE datasets in terms of Dice Coefficient

C to SPIM Assembly Compiler *Spring 2021*
Prof. Uday Khedker | Implementation of Programming Languages *Course Project*

- Implemented a Multi Stage Compiler in C that generates and Abstract Syntax Tree (AST), Three Address Code (TAC), Register Transfer Language (RTL) and assembly code for a subset of programs

Quantum Computing *Autumn 2020*
Prof. Bernard Menzes | Computer Architecture *Course Project*

- Derived theoretical results and quantum circuits for Quantum Weighted Model Counting, an approach that solves the Boolean SAT problem with a quadratic speed up over classical algorithms
- Implemented a proof-of-concept demo using quantum hardware on IBM Quantum Experience

Technical Skills

Programming	C,C++, Python, MATLAB, Javascript, Java, Solidity, C#, VHDL, Bash, HTML/CSS
Tools & Libraries	PyTorch, Tensorflow, Huggingface Transformers, Kaldi, Qiskit, React, Django, Channels, PostgreSQL, Flask, NodeJS, Azure DevOps, Neo4j, GraphAware

Select Coursework

Machine Learning	Artificial Intelligence and Machine Learning, Medical Image Computing, Foundations of Intelligent and Learning Agents, Automatic Speech Recognition, Learning with Graphs
Computer Science	Computer Networks, Operating Systems, Computer Architecture, Design and Analysis of Algorithms, Database Systems, Implementation of Programming Languages, Automata Theory, Abstractions and Paradigms of Programming
Math	Calculus, Linear Algebra, Differential Equations, Discrete Structures, Numerical Analysis, Games and Information

* to be completed by Apr 2022

Positions of Responsibility

- **Content Editor** | *CSE Research Webiste* *Aug 2021 - May 2022*
  Part of a team interviewing 10 faculty members regarding their experiences and research interests to promote undergraduate research, created a repository of information on IITB research labs. Conducted a session to guide over 100 sophomores on seeking research internships
- **Teaching Assistant** | *Data Structures and Algorithms* *Jan 2021 - May 2021*
  Designed competitive programming problems and test cases on HackerRank, graded papers and held regular help sessions for a class of 200 students
- **Teaching Assistant** | *Logic for Computer Science* *Aug 2020 - Dec 2020*
  Developed an effective course evaluation plan for the remote setting, held regular tutorial sessions, graded and released solutions for problem sets. Class of 80 students.
- **Tutor** | *Angel Express Foundation* *Dec 2019 - Jan 2020*
  Taught secondary school level Math and English to a class of 15 at the AXF Non Profit Organization for underprivileged children in Mumbai. Prepared and graded suitable assignments.

Extra-curricular Activities

- Represented IIT Bombay in swimming and water polo at the 35th Inter IIT Aquatics Meet, secured 7th position in the 200m breaststroke across all IITs *2019*
- Secured second position among all undergraduates at IITB Swimathon, covering a distance of 21.1 km in 12 hours *2019*
- Completed the German Language Course offered by International Relations, IIT Bombay *2018-2019*
- Attended the GRA (Group for Rural Activities) Camp to Jhabua village to understand challenges faced in the technical development of rural areas *2018*
- Certified by the Jawahar Institute of Mountaineering for the Mountain Adventure Camp held at Sanasar, Kashmir *2019*