

# RICHEEK DAS

Final-year Undergraduate, CSE, IIT Bombay, Powai, Mumbai - 400076

🌐 [cse.iitb.ac.in/~richeek/](http://cse.iitb.ac.in/~richeek/) | ✉ [richeek@cse.iitb.ac.in](mailto:richeek@cse.iitb.ac.in) | 🌐 [sudoRicheek](https://github.com/sudoRicheek)

## RESEARCH INTERESTS

---

Machine Intelligence, Compressed Sensing, Statistical Learning, Causal Statistics, Sparse Representation

## EDUCATION

---

### Indian Institute of Technology Bombay

*Bachelor of Technology (with Honors) in Computer Science and Engineering*

MH, India

2019 - 2023 (Expected)

- **Major GPA : 9.55/10.0**

- **Thesis** : Code-switched Text Modelling for Natural Language Understanding (*Prof. Preethi Jyothi*)

### Methodist School, Dankuni

*Intermediate. 98.00% in the Indian School Certificate (ISC) Exam*

WB, India

2019

### Methodist School, Dankuni

*Matriculation. 98.80% in the Indian Certificate of Secondary Education (ICSE) Exam*

WB, India

2017

## PUBLICATIONS, PREPRINTS AND WORKING PAPERS

---

5. **Richeek Das**, Sahasra Ranjan, Shreya Pathak, Neel Gupta, Preethi Jyothi, **MLM Pretraining Techniques for Code-Switched Natural Language Understanding tasks**, *Working Paper*
4. **Richeek Das**, Aaron Jerry Ninan, Adithya Bhaskar, Ajit Rajwade, **Bounds for Data-Dependent Weighted LASSO under Low Variance Multiplicative Gaussian Noise**, *Working Paper*
3. Alex Markham, **Richeek Das**, Moritz Grosse-Wenttrup, **A Distance Covariance-based Kernel for Nonlinear Causal Clustering in Heterogeneous Populations**, *accepted at the CLeaR 2022 (1st conference on Causal Learning and Reasoning)*
2. Alexander Erlei, **Richeek Das**, Lukas Meub, Avishek Anand, Ujwal Gadiraju, **For What It's Worth: Humans Overwrite Their Economic Self-interest to Avoid Bargaining With AI Systems**, *accepted at the ACM CHI 2022 (Conference on Human Factors in Computing Systems)*
1. Ashish Tiwari, **Richeek Das**, Shanmuganathan Raman, **Exploring Deeper Graph Convolutions For Semi-Supervised Node Classification**, *accepted at the IEEE ICASSP 2022 (International Conference on Acoustics, Speech, and Signal Processing)*

## RESEARCH EXPERIENCE AND INTERNSHIPS

---

### Neural Architecture Search for Semantic Segmentation

May 2022 - Jul 2022

*Guide: Takuya Narihira, Hsingying Ho, R&D Tokyo Laboratory*

*Sony AI, Japan*

- Proposed a novel Once-For-All based NAS framework for Semantic Segmentation, with plug-and-play features for building dynamic sub-networks with differing hardware-constraints without further fullnet training
- Adapted the method to DeepLabv3+ and proposed a dynamic decoder observing a 7%-15% increase in mIoU
- Developed an EvolutionFinder algorithm to search DeepLabv3+ subnets with FLOPs and latency constraints

### LASSO for RT-PCR group Testing under Multiplicative Gaussian Noise

Jan 2022 - Present

*Guide: Prof. Ajit Rajwade, Department of Computer Science and Engineering*

*IIT Bombay, Maharashtra*

- Provided interpretability and justification for the use Tapestry Pooling across hospitals for pooled testing
- Derived data-dependent weights for Weighted LASSO under Low-Variance Multiplicative Gaussian Noise with physically realizable simulations on RT-PCR group testing data and Kirkman Triple matrices
- Backed up simulations with complete theoretical performance bounds for Weighted LASSO with data dependent weights and showed that with certain sparsity assumptions, it outperforms Plain LASSO

## Code-Switched Text Modelling for Natural Language Understanding

Jul 2022 - Present

Guide: Prof. Preethi Jyothi, CSALT Lab

Google India + IIT Bombay

- Implemented intelligent masking strategies for MLM pretraining of code-switched models and observed immense improvements in downstream code-switched Question Answering and Sentiment Analysis tasks
- Currently working on a generalised framework to adapt existing language translation models for low-resource code-switched text generation with multiple constraints: formality, politeness, toxicity, semantic similarity, etc

## Feature Gating for Deeper Graph Convolution Networks

Dec 2020 - Jun 2021

Guide: Prof. Shanmuganathan Raman, CVIG Lab

IIT Gandhinagar, Gujarat

- Introduced feature gating and formulated a heuristic to award importance scores to nodes and node features
- Proposed the use of identity mapping, a modified form of residual connection and feature gating to create deep GCN models which tackle oversmoothing and achieve SOTA results for semi-supervised node classification
- Performed ablation studies on GCN, GAT, GFGN, GCNII, GRAND with DropEdge, Residual connections and Gating, visualising node importances in t-SNE and feature scores in colour-maps for the Planetoid datasets

## Dependence Contribution Kernel for Non-linear Causal Clustering

May 2021 - Oct 2021

Guide: Prof. Moritz Grosse-Wentrup, Neuroinformatics Lab

Universität Wien, Austria

- Simulated non-linear causal datasets to establish evaluative bounds on the performance of the distance covariance-based dependence contribution kernel and compare it with standard RBF and Polynomial kernels
- Built a module to imitate the problem of causal structure learning in the setting of heterogeneous populations
- Visualised the dependence contribution map (projected kernel space) of causal structure samples by finding an appropriate dimensionality reduction heuristic for the high-dimensional space of causal ancestral graphs

## Belief Elicitation on the Impact of Algorithmic Decision Making

May 2021 - Oct 2021

Guide: Prof. Ujwal Gadiraju, Delft AI Labs

Technische Universiteit Delft, Netherlands

- Implemented Binarized Scoring Rule based criterion for Belief Elicitation of user behaviour, presumptions and trust on the usage of Decision Support Systems (AI-System) for Algorithmic Bargaining
- Built and deployed a DRF backend, Angular 12 frontend, PostgreSQL DB application coupled with Redis + Celery task management, on a Heroku + GitHub deployment pipeline to host 2700+ crowdsource submissions
- Analysed the incentivized subject beliefs and showed statistically consistent results that responders predicting income maximization for the AI agent overwhelmingly override economic self-interest to avoid the algorithm

## MACHINE LEARNING PROJECTS

### Active Learning using Node Embeddings in Partially Observed Networks

Aut 2021

Prof. Abir De | CS768: Learning With Graphs

IIT Bombay

- Combined the SINE: Scalable Incomplete Network Embedding framework and ALPINE: Active Link Prediction Using Network Embedding to build an end-to-end module for Active Learning in Partially Observed Networks
- Proposed ALNEPON, an algorithm for Active Link Prediction in PONs with Incomplete Node Embeddings

### Compressed Sensing Over Graph Structures

Spring 2021

Prof. Ajit Rajwade | CS754: Advanced Image Processing

IIT Bombay

- Implemented the DICE nod algorithm for finding the top information flow hot spots in social networks using compressive sensing with only end-to-end measurements without full knowledge of network topology
- Verified and elucidated the recovery guarantees of the sparse Betweenness Centrality Vector from the proposed Sensing Matrix using properties of Lossless Bipartite Expander Graphs and the Erdős-Rényi generator

### Video from a Single Exposure Coded Snapshot

Spring 2021

Advanced Image Processing Assignment

IIT Bombay

- Implemented a MATLAB solution for coded aperture compressive temporal imaging to recover a sequence of frames from a single coded-snapshot to achieve temporal gains in video acquisition without spatial compromise
- Performed patch-wise reconstruction using OMP algorithm assuming sparsity in a learned dictionary

### UnPlag - Unsupervised Plagiarism Checker

Aut 2020

Prof. Amitabha Sanyal | CS251: Software Systems Lab

IIT Bombay

- Worked in a team of 3 to build an effective web-based pairwise plagiarism checker for source code files.
- Applied TF-IDF metric on Abstract Syntax Trees for C++ and Python using Clang and ast module
- Integrated the computational model with a multithreaded Django REST Framework backend, an Angular 9 frontend and a NodeJS CLI with a secure file-server and stateless JWT authentication

### X-Ray Anomaly Detection Using CNNs

Summer 2020

Institute Technical Summer Project | Ranked among the Top 3 out of 80 projects

IIT Bombay

- Led a team of 4 to build a Web-App and an Open-API endpoint to automate the process of examining CXRs
- Built a 5 model ensemble which accurately classifies and Grad-CAM localizes up to 5 common thoracic diseases in a heat map overlay with an AUC of 0.915, which is quite close to the present state-of-the-art of 0.94

## SYSTEM DEVELOPMENT PROJECTS

---

- Cluster Monitoring and Alert System** | *Database & Information Systems Course Project* (Spring '22)
- Built a web-app with **DRF**, **Angular** and **InfluxDB** – **telegraf** servers on host machines to perform cluster profiling and alert low-resource warnings in real-time using **socket-servers**, based on thresholds set by users
- Branch Predictors for trace-based Simulators** | *Computer Architecture Course Project* (Autumn '21)
- Implemented branch predictors **TAGE** and **L-TAGE** in **ChampSim**, and performed extensive comparisons with Bi-modal and Hashed Perceptron on the metrics of tag width, MPKI, TAGE Table Size and history length
- Linux Socket Programming** | *Computer Networks Course Project* (Spring '21)
- Implemented the **automatic repeat request** mechanism in C to setup a **fault-tolerant UDP** and compared throughput between TCP variants by implementing experiments with client-server network connections
- Compiler for C-like Language** | *Implementation of Programming Languages Course Project* (Spring '22)
- Constructed a compiler handling a subset of the C language, designing a **recursive descent parser** using **lex** and **yacc** — supports **type inference**, **semantic checks** and translation of **AST** to linear three-address codes
- Extending xv6 Operating System** | *Operating Systems Course Project* (Autumn '21)
- Extended the xv6 OS with syscalls for **demand paged memory** allocation and custom **fork** implementations
  - Implemented **thread synchronization**, **semaphores** and a simple linux-based disk emulated **filesystem** in C

## SCHOLASTIC ACHIEVEMENTS

---

- Among **top 13 out of 1148** students to be awarded a Branch Change to the department of **CSE**. (2020)
- Received an **Advanced Performer(AP)** grade for exceptional performance in Calculus(**MA105**). (2019)
- Secured an **AIR of 497** in JEE Main and **544** in JEE Advanced among **1.2 million** candidates. (2019)
- Secured a perfect **10.0** Semester Performance Index(**SPI**) in Autumn semester of first-year. (2019)
- Received the **INSPIRE** scholarship, awarded to **top 1%** of the **80k+** students in the **ISC** Exam. (2019)
- Received **KVPY Fellowship** for securing **AIR 77** out of **50k+** candidates nationwide. (2018)
- Attained **All India Rank(AIR) of 4** and a **State Rank of 2** in **ICSE** out of **180k+** candidates. (2017)
- Received the **Mamraj Agarwal Rashtriya Puraskar** from the **Governor of West Bengal**. (2017)

## TEACHING

---

- Master TA** | *Department of Computer Science and Engineering* Oct 2022 - Present
- One of the two CSE undergrads selected to introduce TAship to the incoming CSE masters students
  - Responsible for hosting the first offering of TA101, a compulsory course for the future Teaching Assistants
- TA, Software Systems Lab (Excellence in CSE TAship Award)** | *Prof. Amitabha Sanyal* Autumn 2021
- Ideated, framed and graded lab assignments to introduce new frameworks to over 175 CSE sophomores
  - Guided over 30 sophomores in building a multi-framework functional course project to completion
  - Awarded for proactively creating engaging course content and being readily accessible to solve students' issues
- TA, Computer Networks Minor** | *Prof. Vinay Ribeiro* Autumn 2022
- Evaluated labs, theory assignments, graded midterms and cleared doubts for a batch of 100+ students
- TA, Computer Programming and Utilization** | *Prof. Kameswari Chebrolu* Spring 2021
- Taught and closely interacted with a batch of 13 first-year students learning coding for the first time
  - Responsible for conducting regular coding course labs, clearing doubts and to evaluate answer sheets
- TA, Engineering Drawing** | *Prof. Atul Sharma* Summer 2021
- Responsible for conducting and evaluating engineering drawing labs for over 170+ first-year students

## TECHNICAL SKILLS

---

Programming Languages	C++, Python, Java, Bash, sed, AWK, Typescript, SQL, VHDL, Prolog, MIPS
Software Skills	MATLAB, OpenCV, PostgreSQL, Django, Angular, L <sup>A</sup> T <sub>E</sub> X, Quartus
ML & DL Skills	Simulink, DRF, Redis, Celery, BS4, AutoCAD, SolidWorks, Wireshark PyTorch, Keras, NNabla, TensorFlow, Pandas, NumPy, Scikit-Learn

## EXTRA CURRICULARS

---

- Completed **80 hours** of **NSS community** work and mentioned for **exemplary volunteering**. (2020)
- Served as a **mentor** in **Institute Technical Summer Project** for a team of **4 freshmen**. (2021)
- Awarded **2nd** position in **Ad-making** division of **Freshiezza 2k19** under **SilverScreen IITB**. (2019)