

# PRATIK KARMAKAR

PhD Student | MSc in Big Data Analytics | BE in Electrical Engineering

@ pkpratikkarmakar@gmail.com  
✉ pratik.karmakar@u.nus.edu    📍 Singapore    🌐 pk-pratik-karmakar    📄 github.com/pratik2358  
☎ +91 6290517961/ +65 86695024    🌐 https://gitlab.com/pkpratikkarmakar

## RESEARCH INTEREST

Data Quality Management & Data Valuation in Databases and Machine Learning, Privacy Preserving Artificial Intelligence, Optimization for Machine Learning and Deep Learning

## PUBLICATIONS

- **Using a Probabilistic Database in an Image Retrieval Application (EDBT 2025) [Paper]**
  - Utilizing probabilistic data generated by YOLOv5 model for better explanation and fine grained reliability of decision making.
  - Using ProvSQL to compute the probabilities and expected Shapley values in an image retrieval system.
- **Expected Shapley-Like Scores of Boolean Functions: Complexity and Applications to Probabilistic Databases (PODS 2024)**
  - Defining expected Shapley-like scores (including Banzhaf index) for queries on the tuples of a probabilistic database.
  - Computing expected Shapley-like scores on probabilistic databases in polynomial time.
  - Implementing the computation in ProvSQL.
- **Marich: A Query-efficient Max-Information Model Extraction Attack using Public Data (NeurIPS 2023)**
  - Model extraction attack with minimum query budget
  - Model extraction using min-max game to minimize loss between target and attacker model and maximize information leakage
  - Use of Active Sampling methods to minimize query budget by sampling high information queries

## EDUCATION

PhD in Computer Science  
**School of Computing, National University of Singapore**  
📅 January 2023 – Present    📍 Singapore  
CGPA 4/5  
MSc in Big Data Analytics  
**Ramakrishna Mission Vivekananda Educational and Research Institute**  
📅 September 2020 – July 2022    📍 Belur, India  
CGPA 9.25/10  
BE in Electrical Engineering  
**Jadavpur University**  
📅 Aug 2014 – June 2018    📍 Kolkata, India  
CGPA 8.58/10

## TECHNICAL SKILLS

- Python
- R
- PostgreSQL
- Neo4j

## GRADUATE COURSEWORK

- Optimization for Machine Learning
- Data Structures and Algorithms
- Probability and Stochastic Process
- Multivariate Statistics
- Machine Learning
- Computer Vision
- Optimization Algorithms
- Online Learning
- Artificial Intelligence
- Deep Learning and Neural Networks II
- Uncertainty Modeling in AI

## TEACHING EXPERIENCE

Teaching Assistant (as a PhD Student) for courses:

- Data Management and Warehousing [BT5110]
- Database Applications Design and Tuning [CS4221]
- Database Technology and Management [IT2002]

## AWARDS

- Research Achievement Award in the Academic Year 2023/2024 by the School of Computing, NUS
- Teaching Fellowship Scheme awarded by the School of Computing in 2024/2025 Semester 1

## WORK EXPERIENCE

Trainee Data Scientist  
**Abzooba India Infotech Pvt. Ltd.**  
📅 July 2022 - Dec 2022    📍 Kolkata, India  
Developed a data imputation framework for general use in the company's business use cases and mentored two interns on projects focused on Explainable AI.

Graduate Engineer Trainee  
**Vedanta Aluminium Ltd.**  
📅 Sept 2018 - Feb 2019    📍 Odisha, India  
Managed a team for operations and maintenance of Ash Handling Plant and Coal Handling Plant of a power plant (Electrical Department)

Summer Intern  
**Defence Research and Development Organization**  
📅 May 2017 - June 2017    📍 Hyderabad, India  
Design of Brushless DC Motors for Tactical Aerospace Vehicle