

SAYAK CHAKRABARTY

pidnas94335@gmail.com ◇ www.linkedin.com/in/sayak-chakrabarty-cs/ ◇ hellokayas.github.io/
Graduating in Dec 25/ Jan 26 (expected)

EDUCATION

Northwestern University - PhD Computer Science	2024-Present
Northwestern University - M.S. Computer Science, GPA - 3.8/4.0	2021-2023
Indian Statistical Institute - Masters of Mathematics	2018 - 2020
Chennai Mathematical Institute - Bsc.(Hons.) Mathematics and Computer Science	2015 - 2018

PUBLICATIONS

NeurIPS 2023 [Single-Pass Pivot Algorithm for Correlation Clustering. Keep it simple!](#)
NeurIPS 2023 [On The Consistency of MLE Of Probabilistic PCA](#)
IEEE DASC 2022 [Dynamically Changing Attack On Review Fraud Systems \(BEST PAPER AWARD\)](#)
IEEE TCSS Journal [Dynamically Adaptive Defense of Review Fraud Detection Engines](#)
(Submitted) ReadMeReady: Fine-Tuning the Art of Free Documentation

INDUSTRY EXPERIENCE

Nokia Bell Labs ML Research Intern-2024 My research is on exploring various directions in Personalized Federated Learning.(Ongoing project)

PROJECTS

ReadMeReady: Fine-Tuning the Art of Free Documentation

- Designed Large Language Model(LLM) based open source application that developers can use as a support tool for generating documentation of any open source repository for free
- The application offers the user a choice of fine-tuning the model using LoRa as the cost of the user's GPU. The training dataset that we created is publicly available.

[Dynamically Changing Attack and Ensemble Defense Regarding Review Fraud Systems](#)

- Collaborated within a team to successfully implement a **reinforcement learning environment** to detect **Fraud accounts and fake reviews**
- Contributed to the design and implementation of a **novel tree structure specifically engineered to manage extensive review datasets** from e-commerce platforms. This required **mathematical foundations of statistics, machine learning, and optimization.**

US Patent(Submitted): Judicial Support Tool for Finding the K-Most Likely Judicial Worlds

- Collected video data from YouTube and US court websites and **annotated and segmented them to train ML models** for predicting the results of a court case. This was done through documentation and conceptualization and required collaborative teamwork.
- Designed a **customized hit-and-run sampling methodology** and **Logic programming/ML techniques** were applied after **exploratory data analysis, statistical analysis, testing, and model development.**

[Prophet Algorithm For Stock Price Prediction](#)

- **Extracted** stock data sourced from Yahoo Finance.
- Applied predictive modeling and in-depth analysis coupled with data-driven approach captured **trends within the stock market** by using the **Prophet algorithm** and other **linear models, multivariate analysis, and sampling methods.**

SKILLS

Languages:	Python, C++, PHP, HTML, R, Haskell, SQL,Java
Frameworks:	Numpy, Pandas, Scikit-Learn, PyTorch, Matplotlib, Seaborn, Networkx,
Software Tools:	Git, Excel, LaTeX, A/B tests, statistical experiments
ML Techniques:	Supervised/unsupervised/Reinforcement Learning, Few-shot Learning, Transformer Models

OTHER ACTIVITIES

- Served as a reviewer for European Symposium on Algorithms(ESA 2024)
- Served as Teaching Assistant for *CS 336- Design and Analysis of Algorithms* and *CS 212- Foundations of Computer Science*