







SAYAK CHAKRABARTY

 [linkedin.com/in/sayak-chakrabarty-cs](https://www.linkedin.com/in/sayak-chakrabarty-cs)  <https://www.github.com/hellokayasgithub.com/hellokayas>
 +1-802-698-3809  pidnas94335@gmail.com
 2233 Tech Dr, Evanston, IL 60208  <https://hellokayas.github.io/>

I work in Machine Learning Theory and Algorithms. I have previously worked on some approximation algorithm problems like **"Single-Pass Pivot Algorithm for Correlation Clustering. Keep it simple!" (Neurips 2023)** and **"On the Consistency of Maximum Likelihood Estimation of Probabilistic Principal Component Analysis"(Neurips 2023)**

Education

-
- | | |
|--------------|---|
| 2022-Present | Graduate Research Assistant, continuing Masters projects, Northwestern University |
| 2021-2022 | Masters in Computer Science, Northwestern University |
| 2018- 2020 | Masters of Mathematics, Indian Statistical Institute (Kolkata) |
| 2015- 2018 | Bachelors of Science(Hons) in Mathematics and Computer Science, Chennai Mathematical Institute |

Professional Experience

-
- | | |
|---------------------|--|
| September 2022 | Judicial Support Tool : Finding the k-Most Likely Judicial Worlds, US Patent : Judicial Support Tool, (Submitted) <ul style="list-style-type: none">> I implemented the whole code base.> I implemented the Hit and Run sampling technique to fit our project.> I collected the whole data used for the project by annotating several hours of YouTube video and breaking them in logical propositions> I ran logic programming techniques on data to generate best worlds <div>Python Pandas Scikit-learn Networkx NumPy Matplotlib Git</div> |
| September 2021 | (Best Paper Award), A New Dynamically Changing Attack on Review Fraud Systems and a Dynamically Changing Ensemble Defense, IEEE Intl Conf on Dependable, Autonomic and Secure Computing |
| September 2022 | <ul style="list-style-type: none">> I worked with Rui Liu(Meta) to implement the reinforcement learning environment> I implemented the tree structure so that a massive review base from e-commerce websites can be efficiently handled. <div>Python Pandas Scikit-learn Networkx NumPy Matplotlib Git PyTorch</div> |
| Jan 2023
Present | Ongoing Work, BOTFARM project, to be submitted <ul style="list-style-type: none">> I collected a huge amount of tweets using tweepy from Twitter which are based on a few fixed topics.> I worked with another student to setup a platform where people will be shown these tweets and they will be given a survey to say their opinions.> I designed a twitter bot detection algorithm completely from scratch that was trained on the data I collected. <div>Python Pandas Scikit-learn Networkx NumPy Matplotlib Git PyTorch</div> |
| Jan 2023 | Fun project, A fast SVD algorithm : SVD is a very common tool but when data size is large, it performs poorly. However, this method goes around the usual mathematical formulation and still gives quite good results!, <ul style="list-style-type: none">> https://github.com/hellokayas/Some-Programming-Samples/blob/master/faster_SVD.py <div>Python Pandas Scikit-learn Networkx NumPy Matplotlib Git PyTorch</div> |
| Jan 2023 | Fun project, Prophet algorithm for stock price prediction, <ul style="list-style-type: none">> I collected stock data by scraping them onto an excel sheet> I used the Facebook Prophet algorithm to analyze the trend and made predictions> https://github.com/hellokayas/Some-Programming-Samples/blob/master/Stock.ipynb <div>Python Pandas Scikit-learn Networkx NumPy Matplotlib Git PyTorch Excel</div> |
| Jan 2023 | Fun project, Participated in "Predict Future Sales" Kaggle competition, <ul style="list-style-type: none">> https://www.kaggle.com/code/saychak/notebook5ead2fa2b7 <div>Python Pandas Scikit-learn Networkx NumPy Matplotlib Git</div> |

Publications

- 2023 | **Single-Pass Pivot Algorithm for Correlation Clustering. Keep it simple!, (Accepted),**
Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS,2023)
[Sayak Chakrabarty](#) [Konstantin Makarychev](#)
- 2023 | **On the Consistency of Maximum Likelihood Estimation of Probabilistic Principal Component Analysis, (Accepted),**
Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS,2023)
[Arghya Datta](#) [Sayak Chakrabarty](#)
- 2023 | **JUST : Judicial Support Tool, US Patent,**
Submitted
[Maksim Bolonkin](#) [Sayak Chakrabarty](#) [Cristian Molinaro](#) [V.S. Subrahmanian](#)
- 2023 | **Judicial Support Tool : Finding the k-Most Probable Judicial Worlds, ,**
Submitted
[Maksim Bolonkin](#) [Sayak Chakrabarty](#) [Cristian Molinaro](#) [V.S. Subrahmanian](#)
- 2023 | **SockDef : A Dynamically Adaptive Defense to a Novel Attack on Review Fraud Detection Engines, (Accepted),**
IEEE Transactions on Computational Social Systems
[Youzhi Zhang](#) [Sayak Chakrabarty](#) [Rui Liu](#) [Andrea Pugliese](#) [V.S. Subrahmanian](#)
- 2021 | **A New Dynamically Changing Attack on Review Fraud Systems and a Dynamically Changing Ensemble Defense, (Accepted), IEEE Best Paper Award**
IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)
[Youzhi Zhang](#) [Sayak Chakrabarty](#) [Rui Liu](#) [Andrea Pugliese](#) [V.S. Subrahmanian](#)
- 2017 | **The Repeated Divisor Function and Possible Correlation with Highly Composite Numbers, Accepted,**
International Workshop for Young Mathematicians "Number Theory"
[Sayak Chakrabarty](#) [Arghya Datta](#)

Skill

Programming Language : **Python**, C++, SQL, Haskell
Operating Systems : **Windows**,Linux
Tools and Framework : **Pandas, Scikit-learn, Networkx,NumPy,Matplotlib,Seaborn**, PyTorch,R, LATEX, Git,Excel
Statistical Skills : **Regression Analysis,Testing of Hypothesis : A/B testing**, Probability theory

Relevant Courses

I have done several courses on Machine Learning, Graduate algorithms, Game Theory, Mechanism Design, Combinatorial Optimization, Advanced Graphics Seminar, Logic in AI. I have TA courses like Analysis II and Foundations of CS.

Awards

- 2022 | **Best Paper Award**, The 20th IEEE International Conference on Dependable, Autonomic and Secure Computing (DASC 2022)
- 2019 | **International Youth Math Challenge**, Indian Institute of Technology
- 2017 | **Science and Engineering Research Board International Travel Award**, Department of Science and Technology, India
- 2014 | **Mathematical Talent Reward**, Indian Statistical Institute