Prateek Chanda

Research Software Engineer (SCAI Center Fellow), Microsoft Research Lab, India
prateekiiest.github.io | Google Scholar | ResearchGate | GitHub | LinkedIn
prateekkol21@gmail.com | +91-8337055526

Experience

Microsoft Research — **Technology for Emerging Markets** — Bangalore , India
Research Software Engineer, SCAI Center Fellow — Supervisor : Dr. Amit Sharma — Nov 2019 - Present

MindNotes - Application for mental health diagnosis

- Collaborating with NIMAHNs and ICARUS UI Design team to build an end-to-end internet based mental health application.
- Implemented an initial architecture for the app using Azure Bot Framework and Azure Cognitive Services - LUIS to understand patient query and showcase relevant mental health resources within the app.
- Developed and shipped entire back end on Azure SQL Server using App Service for storing user responses and client/therapist details and other logistics.
- Integrated Microsoft Graph API with existing App Service for building a message communication framework between patients & therapists with support for audio and note sharing.
- Build NLP models Azure Cognitive services to build word embeddings from free text responses for detecting suicidal responses or sentiments for improving recommendations.
- Content Recommendation via causal inference learning and traditional matrix factorization methods.
- Design effective interventions including content/section recommendation service through causal inference analysis.

Let's Talk: Microsoft Teams App for Mental Health

- Worked in a team of 15 people and led the development group in building an end-to-end Microsoft teams application for recommending appropriate mental health resources within the workplace.
- Built the entire system using Azure Bot Framework to represent a chat framework where users can express their current emotional status and get recommended to appropriate HR resources within Microsoft to cater to their current emotional status.

Indian Statistical Institute — **Indian Statistical Institute** — Baranagar , India
Research Assistant (Remote) — Advisor : Malay Bhattacharyya — May 2021 - Present

- Performed EHR analysis based on Sage Bio Networks Competition to detect which patients are most Covid vulnerable. Employed models like Graph Neural Networks, InterpretML to classify patients and other ensemble classification algorithms
- Proposed a distributed anomaly detection frame work in large data streams in an online manner. Currently under review at VLDB

Complex Network Research Group — **IIT Kharagpur** — Kharagpur , India
Research Consultant (MHRD sponsored) — ***AI for Systems*** — Advisor : Bivas Mitra , Niloy Ganguly — May 2019 - Nov 2019

- Performed a theoretical study on state of the art anomaly detection algorithms in enterprise systems and did a comparative study based on models using GANs and simple classication approaches
- Proposed a data-driven framework for real time anomaly detection/troubleshooting of large-scale storage system failures leveraging several data mining techniques on system logs like regression, clustering, Apriori, graph analytics **See project for details**

Research Intern (Microsoft Research Sponsored) — ***Multi-User Activity Recognition*** — Advisor : Bivas Mitra — May 2017 - Jun 2017

- Performed a theoretical study on traditional group activity recognition models.
- Based on temporal sensor data distribution, estimate missing data through Expected maximization algorithm
- Estimated Group formation over temporal sensor data like SSID-WiFI values
- Studied the correlation effect and casuality for different user features like GPA,locationData w.r.t group activity and formation
- Contributed to the theoretical analysis of the model (initial working of GroupSense) that got acknowledged in Paper. *Project Docs | Code*

Machine Intelligence Unit — **Indian Statistical Institute India** — Baranagar , India
Research Intern May '18 — Advisor : Dr. Ashish Ghosh — Project Report | Code

- Performed theoretical study on different Metric Learning algorithms to learn similarity metric from data distribution.
- Did an empirical analysis as well as evaluation of metric learning methodologies w.r.t different datasets like Iris, Wine Dataset, thus showcasing performance & limitations across various data distribution.

SunPy — **NASA Open Source Software | Solar Data Analysis in Python** — Remote, US
Google Summer of Code — Technologies : Python , Git — Dec 2016 - Apr 2018

- Collaborated with a team of 60 researchers to develop modules for efficient solar data retrieval, data processing and storage functionality for data analysis.
- Implemented a solar data retrieval system to collect solar data from solar observatories based on date in an SQL data base toanalyze different helio-features from the data over a period of 10 years. Used by the SunPy project.
- Implemented proposed solar image processing algorithm from research paper achieving 18% improved memory utilization and better feature extraction with less noise. Got acknowledged along with researchers at NASA Goddard Space Flight Center in nine Software Releases.

Publications

- **MINDNOTES: A Mobile Platform to enable users to break stigma around mental health and connect with therapists**: CSCW Demo 2021
  ***Prateek Chanda**, Amogh Wagh, Jemimah A Johnson, Swaraj Renghe, Vageesh Chandramouli, George Mathews, Sapna Behar, Poornima Bhola, Girish Rao, Paulomi Sudhir, TK Srikanth, Amit Sharma, Seema Mehrotra*
- **Does the Relationship Between Modules Facilitate in Predicting System Anomaly?**: Under Review PAKDD 2022
  *Harsh Borse, **Prateek Chanda**, Paromita Dutta Soumik Ghosh, Mainack Mondal, Bivas Mitra*
- **Human Computation and Crowdsourcing for Earth**: Accepted AAAI HCOMP
  *Yasaman Rohanifar, Syed Ishtiaque Ahmed, Sharifa Sultana, **Prateek Chanda**, Malay Bhattacharyya*
- **Reaching out : Towards a sustainable allocation strategy between users and therapists**: Accepted NeurIPS Machine Learning in Public Health
  Prateek Chanda
- **Distributed Anomaly Detection in Edge Streams using Frequency based Sketch Datastructures**: Under review VLDB 2022
  ***Prateek Chanda**, Malay Bhattacharyya*
- **A Sketch Based Game Theoretic Approach to Detect Anomalous Dense Sub-Communities in Large Data Streams**: Under review AISTATS 2022
  ***Prateek Chanda**, Aadirupa Saha*
- **A Novel Graph Based Clustering Approach to Document Topic Modeling**: Accepted *9th ICCCNT 2018, IISc*
  ***Prateek Chanda**, Asit Kr Das*
- **SunPy A Python package for Solar Physics**: Journal Paper *Journal of Open Source Software 2020*
  *Stuart Mumford, **Prateek Chanda**, The SunPy Community*
- **SunPy v1. 0, the community-developed, free and open-source solar data analysis environment for Python.** : Journal Paper *American Geophysical Unit / NASA ADS 2019*
  *Stuart Mumford, **Prateek Chanda**, The SunPy Community*
- **The sunpy project: Open source development and status of the version 1.0 core package**: *The Astrophysical Journal 2020, IOP*
  *Stuart Mumford, **Prateek Chanda**, The SunPy Community*

Projects

More projects on GitHub : prateekiiest/repositories

Graph Based Clustering Document Topic Modelling
- Designed a novel clustering algorithm based on importance factor calculation of nodes in complex networks with improved accuracy compared to traditional graph based methods. - 9th ICCCNT 2018, IISc, IEEE

Real Time Anomaly Detection in Enterprise environment - NetApp Collaboration ***Advisor: Bivas Mitra, Niloy Ganguly***
Proposed a novel anomaly detection along with failure prediction in an enterprise setting with modules/microservices interacting with each other. Studied how module/microservice interaction changes w.r.t time within the normal period and the anomaly period. For online phase, based on logging intervals proposed a thresholding mechanism to indicate if the corresponding next logging interval is anomalous based on previous behaviour. Presented our method to researchers at NetApp for different enterprise datasets like Hadoop, IBM Bluegene. **Check publications**

Future Promising Heavy Hitter Detection in Streaming environments Advisor : Bryan Hooi, NUS (2020 - July 2021)
- Proposed a novel anomaly scoring technique for the application of heavy hitters using apache datasketches frequent sketches and cumulative distribution comparisons based on quantile sketches. - Performed statistical tests to validate the accuracy of sketches for large dataset items in our anomaly scoring function via sketch guarantees and confidence bound. Currently work in progress

Recommendations for Mental Health Therapeutic activities in MindNotes - Microsoft Research India Advisor : Dr. Amit Sharma
- Developed user feature embeddings based on user responses logging and telemetry logging for building recommendation models for recommending different therapeutic activities and sections of the app as interventions to the user.
- Started with initial collaborative and content based filtering for recommendations, and further developed a causal recommendation model where each micro-intervention was indicated as treatment to the user.

Achievements

- **AISTATS 2022 Mentorship Program**: Selected for AISTATS Mentorship Program to work with Aadirupa Saha at Microsoft Research NYC
- **AI for Science NeurIPS Workshop Mentorship Program**: Selected for AI for Science NeurIPS Mentorship Program to work with Malay Bhattacharyya at Indian Statistical Institute
- **Microsoft Garage Hackathon 2020**: Recipient of Hackathon 2020 NGO award from Microsoft Garage India under AI for Social Good.
- **GAABESU research award IIEST**: Received GAABESU(IIEST) research award for research contributions for academic year 2018
- **JBNSTS Scholar**: Jagadis Bose National Science Talent Search - 2013
- **RMO**: Qualified for Regional Mathematics Olympiad
- **AIEEE Merit**: Within top 0.26% of applicants in All India Engineering Entrance Exam approx. 1.3 million people
- Qualified for **Facebook HackerCup 2019** & **Google Kickstart 2019**

Skills

- **Languages**: Python, C++, SQL, Java, C#, TypeScript **Technologies**: Azure, Azure ML Studio, GitHub, GitLab, Jekyll, GCP
- **Libraries**: TensorFlow, PyTorch, Scikit-Learn, Pandas, Jupyter, Microsoft Graph SDK

Education

Indian Institute of Engineering Science and Technology, Shibpur Howrah, India
Bachelor of Technology in Computer Science & Engineering; *First Class Honors* GPA: 8.86/10.0 WES: 10/10 2015 - 2019

Thesis: *Avoiding Past Choice Regrets: A Game Theoretic Community Detection using Temporal Information*
Advisor: Malay Kule & Dr.Susanta Chakraborty Thesis Report

Relevant Coursework: Data Structures & Algorithms, Operating Systems, Database Management, Cloud Computing & Big Data, Machine Learning & AI, Probability & Statistics, Discrete Structures, Computer Graphics,Computer Networks, Computer Architecture

Positions of Responsibility

- **Program Committee for ODD SIGKDD workshop** 2021 - Workshop Link
- **Reviewer** for IEEE Transactions on Mobile Computing , COMSNETS, AISTATS, Journal of Open Source Software
- **Google Code In**, **GSoC Mentor** : Mentored over 80 students under Google Code In 2018, Hacktoberfest 2018, 2017
- Leading the open source club at Campus as a GitHub Campus Expert - organising hackathons and open source mentorship programs in campus and engaged students from different departments in open source