Mayukh Bhattacharyya

mayukhbhattacharyya18@gmail.com | +1 631-977-9824 http://mayukh18.github.io

EXPERIENCE

Snapchat, Machine Learning Engineer Intern — May 2022 - Aug 2022, Los Angeles

• Worked in the Lens Ranking team on the development of offline evaluation methods for large-scale recommendation systems. Developed counterfactual evaluation methods that can closely emulate online A/B test results from historical data, thus reducing reliance on A/B tests for model evaluation.

• Carried out extensive analysis and experimentations involving the developed methods on historical data using GCP and BigQuery, which showed its superior efficiency over traditional metrics. Implemented and integrated the offline methods for the in-house ML platform using Apache Beam.

Stony Brook Biomedical Informatics, Senior Research Aide — Mar 2022 - Present, Stony Brook

• As a part of the National COVID Cohort Collective, analyzing large-scale patient data to uncover patterns involving different aspects of long-COVID.

• Work involves handling terabytes worth of data using Spark, visualizing and drawing conclusions, and presenting the findings to multiple stakeholders.

Applied Materials, Algorithm Developer - II — May 2020 - Aug 2021, Bangalore

• Worked in the CTO research team on the development of an AI-based product for cancer diagnosis using microscopy tissue images, with a particular focus on explainable AI. Built the algorithmic core of the product from ideation to the production-ready stage in under 12 months within a small team.

• Developed the core Computer Vision models for tumor segmentation and cell detection with extremely scarce training data leveraging weakly supervised techniques. Designed and built efficient latency-focused data pipelines for running the model inference in production on gigapixel-sized images.

Sigtuple Technologies, Software Engineer — Oct 2019 - Mar 2020, Bangalore

• Worked in the Data Science team on the development of object detection algorithms for AI-based pathology tests. Developed YOLO-based models which improved detection precision by 30%. Developed algorithms for a pathology imaging device, interfacing device hardware with scanner cameras.

Samsung Electronics, Software Engineer — Jul 2017 - Oct 2019, Delhi NCR

• Worked in the Enterprise Displays Team, where I developed multiple core products in C/C++ with complete ownership for large format displays - video content player, weather app, network-based multi-tv video wall, synchronized play etc. for a Linux-based OS framework.

• Led the research and development of gaze estimation and object detection-based content/advertisement recommendation system for large outdoor displays. Used crowd gaze as an implicit feedback for suggesting better contents to be played on screen.

Goldman Sachs, Technology Analyst Intern — May 2016 - Jul 2016, Bangalore

• Worked in the Finance Data Engineering team. Developed a fault diagnostic safety layer for a Java-based production data pipeline. Developed visualization dashboard in AngularJS for monitoring the overall status of various data flows consolidating the ledgers of the finance division.

PUBLICATIONS

- DeCAtt: Efficient Vision Transformers with Decorrelated Attention Heads — **ECV Workshop, CVPR 2023**
- SERF: Towards better training of deep neural networks using log-Softplus ERror activation Function. — **WACV 2023** (link)
- Deciphering Environmental Air Pollution with Large Scale City Data. — **IJCAI 2022** (link)
- Hybrid Style Siamese Network: Incorporating style loss in complementary apparels retrieval. — **CVFAD Workshop, *CVPR 2020*** (link)

EDUCATION

Stony Brook University, M.S. in Computer Science — Aug 2021 - May 2023*, NY

Jadavpur University, B.E. in Electrical Engineering — Aug 2013 - Jun 2017, Kolkata

PROJECTS

reco — Recommender Systems library (link)

Created a comprehensive Python library of Recommender Systems with core modules of Collaborative Filtering, SVD, Factorization Machines, Wide and Deep Networks, etc. with an underlying cython core. The published package has been downloaded over 13k times on PyPI.

BlindChat— Flask based chat app on FB Messenger (link)

Built and deployed a cross-matching chat app in Flask on top of Facebook's messenger APIs. Implemented server backend with user data stored in SQL database. It had over 1/2 million messages exchanged in 2017.

CovidExplore — exploring impacts of COVID (link)

Developed a website that demonstrates the impacts of COVID through interactive visualizations based on bokeh. Scraped the data from various online resources reflecting the impact of COVID on different aspects of the world and society.

SKILLS

Python, C++, Java, Javascript, SQL, Machine Learning, Deep Learning, Computer Vision, Operating Systems, Data Structures, Algorithms, Recommendation Systems, Big Data, NLP, Linux, Flask, PyTorch, Tensorflow, Numpy, OpenCV, Scipy, Spark, Apache Beam, GCP, Git