RAJ RAY CHAUDHURY

DREXEL UNIVERSITY College of Computing & Informatics

115 North 34th Street, Philadelphia, Pennsylvania 19104

📞 215-730-3924 ✉ arr349@drexel.edu ✎ https://rajrayy.github.io/

Education

Drexel University, Pennoni Honors College   **Cumulative GPA: 3.93**
*Bachelor of Science in **Computer Science (Honors)**, <u>Minors</u> in **Math** and **Data Science***   *September 2019 - Present*
- **Dean's List**(All quarters)
- AJ Drexel Merit Scholarship (2019 - present)
- Clubs - i) Drexel Consulting Group **(Board Member)** ii) Drexel Algorithms and Data Structures (Member)

Experience

College of Computing and Informatics   **Apr 2022 - Present**
Teaching Assistant for Advanced Python (CS - 172) & Introduction to Bash and C (CS 265)   *Philadelphia, Pennsylvania*
- Walked students through weekly labs and clarified concepts on fundamental data structures like stacks, queues, hash maps etc.
- Held weekly office hours, provided continuous feedback and assisted the instructor in compiling coursework for over **150** students.
- Utilized **auto-grading** script that uses regular expressions to transfer grades from third party ed-tech software to student portal.

Amazon.com, Inc.   **Sept 2022 - Mar 2023**
Software Development Co-Op   *Boston, Massachusetts*
- Employed within the **Robotic Perception and Manipulation team** where I curated **ARMBench**, a large-scale benchmark dataset **(250k+ images** and **4000+ videos)** with **high quality annotations** collected in an Amazon warehouse.
- Developed the entire software architecture from scratch for an **automated data ingestion pipeline** that picks raw data from **S3** and ultimately submits clean data into the benchmark dataset.
- Wrote **multithreaded** scripts to implement pre-defined **sampling criteria** for picking items for annotation from raw station data.
- Built a custom **AWS CDK** package that **listens** to event triggers, fires off **Lambda handlers** and controls an **EC2** instance.
  - Lambda A runs scripts that extracts relevant images and **contours segmentation masks** using **opencv** and **numpy**.
  - Lambda B **crops regions of interest** based on use cases and **stitches 4 viewpoint** images together for annotation.
  - Lambda C queries an in-house **item matrix** database to identify objects in images before deploying a **multi-label image classification job** on **AWS Sagemaker**.
  - Lambda D parses through **batch annotations** and uploads relevant data attributes to a **DynamoDB** database.
  - Designed **worker task template** to define the UI, **workforces** and **generate ground truth files** with human annotations.
- Used **PartiQL** to loop through DynamoDB to **retrieve relevant images and videos** before pushing to the cleaned dataset.
- Developed a **static website** that allows external users to **download the dataset** and keep up to date with progress. 🔗

Susquehanna International Group LLP   **Sept 2021 - Mar 2022**
Software Development Co-Op   *Bala Cynwyd, Pennsylvania*
- Employed within the **Electronic Options Trading team** where I assisted the development, delivery and support of technical solutions for various proprietary and third party applications systems.
- Developed a **.NET app** that uses **12** parameters to audit up to **500** log files **at a time** through simultaneous **optimized parsing**.
  - Leveraged Python's **multiprocessing** module to maintain shared file queues and improve back-end processing time by **55%**.
  - Maintained fully **automated CI/CD** pipelines for code deployment using **GitLab, Team City and Octopus Deploy**.
- Mitigated production incidents and saved **technology dollars** through a script (written from scratch) that automatically **detects** and **flags duplicate IP addresses** where both host and alias may have jobs scheduled in **Tidal** (automation software).
- Built a **REST API** in **Flask** & **Postman** that lets end users generate **CSVs** by querying a range of historical **MySQL** databases.
- Generated a **dashboard** that allowed my manager to assign/remove read/edit permissions for viewing **sensitive reference data** on a case-by-case basis (**C#** front-end that interacts with an **Oracle** database to pull employee records).
- Built a **chrome extension** in **JavaScript** that simplifies developer workflow and displays crucial server **metadata** with a click.

STAR Research Scholar   **Jan 2021 - Sept 2021**
Drexel University, Prof. Weimao Ke   *Philadelphia, Pennsylvania*
- Implemented a new algorithm (based on existing TF*IDF vectorizer) for **automated test classification** in **Jupyter Notebook**.
- **Cleaned test data** by grouping inflected forms of words through the **Wordnet Lemmatizer** in Python's **NLTK** library.
- Removed stop-words with an **iterated Lovins stemmer** through **tokenization and normalization** of documents.
- Used **k-NN (k-Nearest Neighbours)** as the core algorithm to **train and test both models** on **10000** randomly selected articles from the New York Times annotated collection.

Drexel Solutions Institute   **Mar 2021 - Sept 2021**
*Consultant for **Longwood Gardens***   *Philadelphia, Pennsylvania*
- Identified **research designs** to integrate **VR/AR technology** into one of the world's **premier** horticulture gardens.
- Conducted thorough **literature review**, designed studies using **conjoint analysis** and **analyzed survey data** from **Qualtrics**.
- **Scraped** and **grouped** over **4700** Google reviews to determine visitor location preferences prior to digitization.
- Examined over **450** sets of **eye-tracking** (OGAMA) and **EDA** (electrodermal activity) data in **R** to aggregate clean data.
  - Developed and deployed (via **Heroku**) an analytics app that demonstrated impact of sound complexity on consumer behavior.
- **Visualized** collected data and **presented** solutions to an interdisciplinary audience of industry and academic stakeholders.

Sharing Excess Inc.   **Sept 2020 – Jan 2021**
Product Management Intern   *Philadelphia, Pennsylvania*
- Worked in a team to develop a **full stack web app** in **Python's Flask** framework that **saved 50 man hours** a month by **automating the process of scheduling** food deliveries between donors and recipients.
- Implemented **logging and tracking features** and stored relevant data in a **PostgreSQL database** as an incentive to all stakeholders to achieve our target of **100000 pounds** of food delivered **in 6 months.**

Projects

Private Wealth Manager | *Python*   **Dec 2022**
- Applied the **Sortino Ratio** formula to calculate and compare the profitability and risk of different investments using pandas.

Data Parser - CS 265 | *Bash, AWK, Linux*   **Apr 2021**
- **Bash script** that **generates XML files** to display a **subset** of US Army veteran data on **formatted webpages.**
- Used **AWK** to scrape relevant data from a **temporarily created CSV** file for every legitimate combination of input.

Broke Bad | *React.js*   **Jun 2019**
- **Front-end web application** to display **randomly generated** Breaking Bad quotes pulling data from an existing API.

Technical Skills

Languages: Python, Java, JavaScript, C, C#, C++, HTML/CSS, JavaScript, SQL, AWK, Bash, Racket, LaTeX, Microsoft Excel
Tools — Operating systems: .NET, Heroku, JQuery, React.js, Bootstrap, GitHub, Jira, Tableau, Qualtrics — Linux, Windows