Saurabh Mahesh Raut

sraut@seas.upenn.edu | (+1) 267-902-6290 | linkedin.com/in/saurabhraut99

Education

University of Pennsylvania (UPenn) ***Aug 2021 - May 2023***
Master of Science and Engineering in Computer and Information Science *CGPA: 3.87/4*

Veermata Jijabai Technological Institute (VJTI) ***Aug 2017 - May 2021***
Bachelor of Technology in Computer and Information Science *CGPA: 8.1/10*

Technical Skills

- **Programming Languages**: Java, C/C++, Python, HTML, CSS, JavaScript, SQL
- **Frameworks & Software:** MongoDB, Neo4j, NodeJS, TestNG, Scikit-learn, OpenCV, Pandas, MapReduce, Apache Spark
- **Development Tools:** CDK, AWS (RDS, DynamoDB, S3, SQS, EC2, Lambda), Android Studio, Git

Relevant Experience

Amazon Music (Software Development Engineer) | *Seattle, USA* ***May 2022 - Aug 2022***
- Designed, implemented and deployed a caching framework for Amazon Music's 3P evaluation framework
- Extensively used CDK and AWS tools like Lambda, S3, SQS and DynamoDB for implementing the cache

Barclays Investment Bank (Software Development Engineer) | *Pune, India* ***Jun 2020 - Jul 2020***
- Automated testing of Barclays trade application - BATMAN by using Barclays Automation Test Harness System(BATHS) and testNG framework

Silicon Veins (Software Development Engineer) | *Mumbai, India* ***May 2019 - Jun 2019***
- Built custom Share View replacing the built-in android chooser for BizAnalyst android app. This feature lets users directly share ledgers to mobile numbers through Whatsapp without saving the number in their contacts
- Implemented passcode activity, floating action button and other UI related features for the app

Selected Projects

Corrupt image restoration (Research - Prof. Insup Lee) | *Python, OpenCV, PyTorch, GIMP* ***Jan 2023 - Apr 2023***
- Contributed to building the Supervisory system for Shift Adaptation and Recovery (SuperStAR) algorithm that uses Reinforcement Learning to reverse some semantic-preserving shifts on images
- Improved upon the actions taken by SuperStAR algorithm by experimenting and analyzing changes in state space of images

Penn Search - Distributed Search Engine | *Java, HTML, JavaScript, AJAX, AWS tools* ***Apr 2022 - May 2022***
- Built a completely distributed search engine consisting of Web Crawler, Indexer, Page Ranker, & an interactive UI
- Crawled more than 1 million web pages using StormLite, a lightweight version of Apache Storm, running on EC2 instances. SQS was used as a crawl frontier and the crawled pages were stored in a S3 bucket
- Indexed and ranked the crawled pages as Hadoop MapReduce tasks and handled search queries from user using self-built version of Apache Spark

Penn Web Framework | *Java* ***Mar 2022 - Apr 2022***
- Built a robust miniature version of Spark Java Framework consisting of a server thread pool and request handlers that accept and process client's HTTP tasks by executing the developer supplied request handler function

Penn Cloud | *C++, HTML, CSS* ***Nov 2021 - Dec 2021***
- Built a cloud platform that supports web-mail service and storage service analogous to Gmail and Google Drive
- Storage was based on distributed key-value store (similar to Bigtable) and factors like load balancing, fault tolerance, replication, consistency were taken into consideration as well

NBA Game Info - Web Application | *SQL, AWS RDS, NodeJS, React, HTML, CSS* ***Nov 2021 - Dec 2021***
- Built a web-application that compiles detailed information about each NBA game and NBA player's profile and displays it based on the users needs

Chat Servers | *C++* ***Oct 2021 - Nov 2021***
- Built a replicated chat server that uses User Datagram Protocol (UDP) to multicast chat messages
- The server supports three different ordering modes: unordered, FIFO, and total ordered multicast

Hateful multimodal content detection | *Python, PyTorch, Numpy* ***Oct 2021 - Nov 2021***
- Built hateful content classifier and used Facebook AI hate meme challenge dataset with benign confounders
- Trained and tested using multiple approaches like early fusion and late fusion using pretrained and fine-tuned Resnet image embeddings and RoBerta sentence embeddings

Content Labelling (Research - Prof. Sunil Bhirud) | *Python, Scikit-learn, Gensim, NetworkX* ***Jan 2021 - Apr 2021***
- Scrapped and performed topic classification of approximately 50 thousand web pages from the darknet
- Used two distinct approaches- Topic Modelling & Graph of Words, computed similarity scores and got an accuracy of 78.05%

Relevant Coursework

Software Systems, Internet and Web Systems, Applied Machine Learning, Principles of Deep Learning, Natural Language Processing, Linear Algebra and Optimization, Database and Information Systems, Analysis of Algorithms, Engineering Negotiation, Engineering Entrepreneurship, Data Structures & Algorithms