Raunaq Jain

(716) 380 8692 | raunaqja@buffalo.edu | www.linkedin.com/in/jainraunaq/ | www.raunaqjain.com | github.com/raunaqjain

EXPERIENCE

Software Development Engineer, Audible 04/2021 – Present

- Led the migration of **950+** hosts to Amazon Linux 2 for enhanced performance and security benefits.
- Developed an Integration Test package for Audible Rights Service which currently serves **500** requests of batches per sec.
- Improved team's agile processes as a scrum master and is part of the on-call rotation.
- **Pipeline Ambassador**: Responsible for monitoring the health of **50+** pipelines and lead them towards full CI/CD.

Researcher, University at Buffalo 06/2020 – 12/2020

- Artificial Intelligence in Ethnobotany (AIE): Pioneered machine learning solutions over SSD, YOLO, and Detectron2 for plant species identification in images (**50k**) with applications in Ethnobotany and results published in **ICMLA** conference.
- Secure Storage for TrustZone: Designed and ported secure file storage system from **C to Rust** for Trusted Execution Environment and provide proof of memory safety through software verification.

Machine Learning Research Intern, Hypothizer Technologies Private Limited 07/2018 – 02/2019

- Developed a cloud-native document parsing platform to extract critical information from unstructured documents.
- Architected, developed, and deployed Resume Parser with average F1 score of **0.93** after training on 150 samples in PyTorch.
- Improved accuracy of the existing document parser by **10%** through feature engineering and customized models.

Software Developer Intern, All India Council for Technical Education 10/2017 – 03/2018

- Created a full-stack web-application in Django to set standards for **14k+** technical colleges throughout India.
- Developed the data pipeline and trained machine-learning models to predict the employment potential of universities using scikit-learn. Project sanctioned by the Government of India [Github].
- Curated, analyzed, and visualized **5 years of data** in Python and produced interactive plots through Plotly for user interface.
- Awarded **1st prize** ($3000) in Smart India Hackathon'17, Government of India, against **7400+ teams** for the prototype.

EDUCATION

- University at Buffalo**,** The State University of New York, Masters in Computer Science [GPA: 3.85/4] 08/2019 – 02/2021
- MAIT, Guru Gobind Singh Indraprastha University, B.Tech. in Computer Science [CPI: 71.98%] 08/2014 – 05/2018

Research Paper

- M Böhlen, **R Jain**, W Sujarwo, V Chandola
  From images in the wild to video-informed image classification, 20th IEEE ICMLA, 2021[pdf]

PROJECTS

FEVER: Fact extraction and Claim verification

- Indexed **5 million +** Wikipedia articles on Apache Solr, thus reducing the document retrieval time by **99%** on average.
- Designed a component in PyTorch to evaluate the contextual similarity between two evidence sentences.
- Developed an application in Flask to verify credibility of an input claim sentence. Extracted evidence from Wikipedia articles.
- Verified credibility of input claims and achieved a FEVER score of **41%** (**27% baseline)** by training on **185K claim samples.**

DogDogGo, Search Engine Application [Demo]

- Implemented an end-to-end tweet search engine and built the application backend in Django (REST API).
- Features developed included tweet ranking, pseudo relevance feedback, more-like-this, filter-based search, query translation, and query keyword highlighting.
- Collected **900,000+** multi-lingual tweets. Preprocessed and indexed them on Apache Solr.
- Assessed impact of political rhetoric in traditional and social media, presented through interactive geospatial plots.

Simple Amazon DynamoDB

- Built a concurrent and available distributed multicast messaging system for 5 nodes with fault tolerance and node recovery using sockets in Java and Android.
- Programmed a distributed key-value storage and used Quorum algorithm for communication between multiple nodes.

Visual Question Answering through Modal Dialogue (VQAMD) - 2013

- Developed an application which answers any open-ended question about an input image through a CNN + LSTM model.
- Attained an accuracy of **54.88%** on version 2 of the Visual Question Answering dataset.

TECHNICAL SKILLS

Languages: Python, Java, Rust, TypeScript, AWS CDK
Libraries: Django, Flask, PyTorch, keras, OpenCV, scikit-learn, SpaCy**,** Plotly, matplotlib, NumPy, Pandas
Databases: SQL, MongoDB, DynamoDB
Others: Git, Docker, Apache Solr, AWS, Azure, Agile