HARIHARAN NAGASUBRAMANIAM

(929)678-7353 | hnnhariharan12@gmail.com | soester10.github.io/portfolio | linkedin.com/in/hariharan-naga/

EDUCATION

New York University   New York, NY
MS Computer Engineering | GPA: 3.89   Aug 2022 - May 2024

SRM University   Chennai, India
B.Tech Computer Science and Engineering | GPA: 7.9/10.0   Jul 2017 - Jun 2021

Coursework: Data Structures, Algorithm Design Analysis, Machine Learning, Computer System Architecture, Deep Learning, High Performance ML, Software Engineering, Internet Protocols, Network Security, Big Data

EXPERIENCE

Secure Systems Lab, New York University   New York, NY
Research Assistant | *CodeQL, AST, Software Engineering*   June 2023 – Present

- Investigating the role of "atoms of confusion" in **software bugs** and issues, conducting **statistical analyses** across multiple open-source repositories.
- Generating **Abstract Syntax Trees** for various codebases, enabling in-depth code structure and logic flow analysis.
- Enhancing existing classifiers using static code analysis tool **CodeQL**, achieving over a **500%** speed improvement, and demonstrating a **20%** increase in accuracy and coverage through comparative analysis.

Dun & Bradstreet   Chennai, India
Data Scientist | *Python, SQL, Spark, Pandas, Selenium, Databricks*   Aug 2021 – July 2022

- Developed and scaled **multiprocessing web crawlers**, automating the extraction of ESG-relevant data for 2.5 million companies, encompassing over **60%** of the company's internet-based data.
- Optimized **PySpark** workflows with a single efficient User Defined Function (UDF), resulting in a **30%** increase in Reason Code production speed, including revised logic.
- Constructed an end-to-end **NLP pipeline** for the extraction of pertinent data from ESG reports and news for more than **200k** companies across 15 global markets.
- Integrated the SIC (Standard Industrial Classification) method into production code to shortlist over **5 million** businesses engaged in controversial activities.
- Developed multiple efficient UDFs to compile Self-Assessment Questionnaires from individual companies, effectively handling over **50%** of new data sources for scoring conversion.

PROJECTS

Show Of Hands | *Django, JS, AWS, Travis CI/CD, PostgreSQL, S3, Redis*

- Implemented a dynamic **social app** for polling and user interaction, promoting engagement and live discussions, with social networking features, including friend requests for user connection.
- Engineered a robust **chat system** with **Redis Cache memory** to facilitate **real-time user communication** and in-depth discussions; integrated group chat functionality for smaller, exclusive discussions.
- Introduced unique features like timed result delays to stimulate increased interaction and live comment sections to facilitate debates.
- Streamlined deployment processes with **AWS Elastic Beanstalk** and **Travis** while employing **PostgreSQL** for scalable management of user profiles and interaction history

Fire Chat | *Flask REST API, React, Firebase, AWS*

- Developed an interactive **discourse platform** tailored for **movie enthusiasts**, fostering engaging conversations within a passionate community.
- Designed secluded rooms with comprehensive movie information imported from **OMDB API**, allowing users to post reviews and engage in networking and one-on-one discussions.
- Implemented **Flask REST** framework for the backend, hosted on **AWS**, and frontend using **React** on **Firebase**; ensured data security and real-time chat functionality with **Firestore Storage** and **Firebase Security Functions**.

BRViT | *PyTorch, Computer Vision, ViT*

- Developed an end-to-end **Vision Transformer** model tailored for Bokeh Rendering, enhancing subject of images.
- Utilized **open-source** DPT model as the foundation for encoder, reassembly, and fusion layers in the architecture.
- Demonstrated exceptional performance by achieving **state-of-the-art** results on the challenging EBB! Dataset.
- Significantly reduced computation time by eliminating the need for a separate depth estimation model.

SKILLS & CERTIFICATIONS

Programming Languages: Python, JavaScript, Matlab, C/C++, TypeScript, Java
Tools & Frameworks: Django, Flask, FastAPI, React, PyTorch, TensorFlow, Docker, Terraform, Kubernetes, Kafka
Databases & Platforms: SQL, PostgreSQL, NoSQL, Spark, MongoDB, AWS, Google Cloud, Azure, Linux/Unix
Certifications: Google IT Automation Certificate (Coursera, Jun21), **Microsoft Certified: Azure Fundamentals** (May21), **TensorFlow Developer Certificate** (Coursera, Sep20)