Samir Satam

ssatam@gmail.com ❖ (978) 930-1109 ❖ 19 Woodlea Rd, North Andover, MA 01845

SKILLS

Machine Learning: LangChain API, OpenAI, Mistral, LLM Fine Tuning, PgVector, Pinecone, ML and Statistics.
Tech Stack: Java, Python, Spring, Spring Boot, Kafka, RabbitMQ, GoLang, GRPC, Webservices, Microservices, Kubernetes, Docker, React, Kafka, Linux, Git, GitLab, Github.
Databases: Postgres, MySQL, MS SQL Server, AgensGraph, MongoDB.
Cloud: AWS, GCP, Pivotal Cloud, Rackspace.
Data Analysis: Python, R
Certification: CompTIA Sec+
Security clearance information provided upon request.

WORK EXPERIENCE

RAFT LLC / KESSELRUN [AIR FORCE]   June 2023 – Present
Principal Engineer / Manager   *Remote / Hanscom Airforce Base*

- As a **Principal engineer** I worked on two separate products for KesselRun.
  - Use LLM to convert USMTF 2000 ATO to Natural Language – This is an experimental project at the AirForce to convert USMTF8 to natural language and back. It uses local instance of Mistral as LLM. PgVector was used to store the embeddings for Similarity retrieval. Efforts to supply accurate USMTF for fine-tuning are ongoing.
  - Rebel Alliance – The middleware that powered the event streaming architecture of KR
    - I worked on stability and performance issues that were causing headaches and crashes in production.
    - This involved optimizing code, writing a recovery mechanism for RabbitMQ queues, and optimizing SQL queries and indexes.
    - I also worked on achieving efforts as the Event store grew too large.
- As a **Manager** at Raft, I am also responsible for managing Senior engineers, collaborating and collecting peer reviews, and guiding them in their career paths.

NOETIC CYBER INC.   Feb 2022 – June 2023
Principal Engineer   *Remote / Boston, MA*

- **End User queries to Cypher queries** - Using LangChain and OpenAI for converting End User queries to Cypher Graph Queries for easier User experience. This helps the End users query the security vulnerabilities in Natural (English) language. This involves Prompt Engineering, implementing RAG using the Schema design from AgensGraph. PgVector was used to store Schema details to enable similarity-based retrieval. Fine Tuning OpenAI LLM with JSONL from the Cypher query training set for more accurate results.
- **Transition from JanusGraph database sto AgensGraph.**
  - Writing a GRPC 'Graph server' that connected AgensGraph to Backend services.
  - Implementing portions of OpenCypher unsupported by AgensGraph as Postgres stored procedures.

OMNIFEDERAL LLC / KESSELRUN [AIR FORCE]   June 2020 – Feb 2022
Senior Software Engineer   *Remote / Hanscom Airforce Base*

  - Carbonite – Mission Planning tool for mission planners to US bases all over the world. This allowed the Air Force to plan missions by showing threats and assets on an interactive map.
  - Triton – Air Tasking Order (ATO) creation tool used for mission planning. Triton consumed events from various Air-Force systems and presented an up-to-date ATO to the Airmen.

SPLITLIMIT STUDIOS, LLC   Apr 2018 – May 2020
Principal Software Engineer / Data Scientist   *Remote / Waltham, MA*

- Leading provider of Pay-As-You-Go insurance software.
  - Secure webservices integrating with BaseCommerce and OneInc digital payment systems.
  - Developed custom integration solutions for NJM Insurance group and Republic Indemnity.
  - Developed new tools for production support and Data analysis reports for business decisions.

COPYRIGHT CLEARANCE CENTER   Apr 2001 – Mar 2018
Principal Software Engineer / Data Scientist   *Danvers, MA*

- Founding engineer working on CCC's Rightslink product suite.

- Contributed directly to Rightslink meteoric growth to 35m annual revenue.

- Lead Engineer implementing Rightslink on major publishers like Spring, Dow Jones, Sage, Wolters Kluwers, BMJ and Wiley to name a few.

- Also responsible for "Republication Licensing" a product that grew over 50% every year.

- K-Means for customer segmentation: CCC helps Publishers understand customer behavior and their demographics. By using the K-Means clustering algorithm, the segmentation tool generates behavior-based customer segments, discovering hidden customers based on their transactions, industry and requester type. These segments are then used by Publishers to target customers to offer better services. The process leverages Scikit-learn and Python to generate clusters and provide valuable Business Intelligence to Product Owners with respect to our Publishers, and to our Publishers a better understanding of their customers.

YANTRA INC   Apr 2000 – Apr 2001
Software Engineer   *Acton, MA*

- Worked with client Interare, Brazil for development of a Web front for ThinkChain, a B2B website for buyers and Vendors in Brazil.

TELTRONICS INC   Aug 1999 – Apr 2000
Software Engineer   *Sarasota, FL*

- Designing and developing a Site-Event-Buffer (SEB), a log collection device for PBX's for easier troubleshooting.

QUARK INC   Apr 1998 – Aug 1999
Software Engineer   *Denver, CO*

- Extensions development for Quark Express, Quark DMS and QPS systems.

CANON INC. / NEWGEN SOFTWARE TECHNOLOGIES   May 1996 – Apr 1998
Software Engineer   *Denver, CO*

- Various projects involving digitization of documents.

EDUCATION

Quantic Business School   2017
Executive Masters of Business Management   *Online*

Mumbai University   1996
Bachelor Of Engineering Computer Science   *Mumbai, India*