Kamen Shah

Mountain View, CA
KamenShah.com
github.com/KamenShah
producthunt.com/@KamenShah

Enterprising software engineer with extensive experience in big data, data science, and machine learning. Bringing a successful track record in technical roles while building novel solutions that are scalable and fault tolerant. Seeking an opportunity to work on AI related applications with a forward-thinking organization.

TECHNICAL SKILLS

Fluent in: Python, Java, SQL, PL/SQL
Experienced with: C++, PHP, HTML5, CSS, Bash, Scala, JavaScript, Golang, Prolog
Tools: Git, Linux, Jira, Spring, Express, Oracle, MySQL, PostgreSQL, MongoDB, Hadoop, Spark, PCF, NGINX, Keras, Kubernetes, Docker, Gradle, Jenkins, PyTorch, TensorFlow, Scikit-learn, Pandas, NumPy, Denodo, Apigee, XGBoost

WORK EXPERIENCE

Software Engineer – Intel (TCS), Remote *12/2020 – Present*

- Spearheading project to accelerate ETL jobs through introducing parallelization, yielding 700% improvement in processing time and saving 1000+ hours yearly
- Conducting PLSQL performance tuning through B-tree indexes, views, partitions, and clusters within an Oracle data warehouse, reducing load time for users in consuming applications
- Collaborating with sales team to reconcile legacy data using fuzzy logic and data transformations with Pandas
- Leveraging Denodo to join disparate data sources and expose it to downstream users through Apigee REST services

Big Data Engineer Intern - Gap Inc., San Francisco *06/2019 – 08/2019*

- Led team of 3 developers to port a monolithic ETL job to the Java-powered Spring Boot/Batch framework, increasing maintainability and extensibility of production code while reducing batch processing time by 25%
- Orchestrated 2-week Agile sprint cycles through Jira and maintained high velocity throughout the entire project
- Designed and automated a Jenkins CI/CD pipeline using Gradle to build and deploy ETL jobs to PCF on Azure

Software Developer/ Co-creator – AMP Ventures, Mountain View *06/2017 – 06/2019*

- Architected full-stack web application that curates live gaming, news, and sports content from the internet to provide an intuitive UI for content browsing, acquiring over 80 DAU and 500 MAU
- Defined and employed a hook model that generated virality through triggers, increasing session duration by 30%
- Scaled application across 4 GCP g1 servers with Docker and Kubernetes, and distributed workload across cluster with NGINX load balancer

SOFTWARE PROJECTS

Master's Thesis: Wildfire Spread Prediction *09/2021 – Present*

- Extracted satellite imagery of wildfire and ancillary events from Google Earth Engine and stored data on GCP
- Designed U-Net architecture augmented with attention layers for image segmentation with TensorFlow

Machine Learning with Financial Data *06/2020 – 08/2021*

- Wrangled 50GB of financial data from online sources using multi-threaded processing over the Tor Network
- Unified, cleaned, and analyzed data with Pandas, NumPy, PCA, encoders, scalers and deep learning imputers
- Built and compared DNNs, LSTMs, attention-based transformers, and ensemble models for stock prediction

Object Detection in Videos with Northrop Grumman (Link) *08/2019 – 05/2020*

- Collaborated with an interdisciplinary team to integrate RNN, LSTM and TCNN layers into a CNN for time-series object detection in PyTorch, identifying greatest improvement in classifications while embedding LSTM layers

EDUCATION

Computer Science M.S. – California Polytechnic State University, San Luis Obispo, GPA: 4.0 *09/2021 – 12/2022*
Computer Science B.S. – University of Colorado Boulder, Major GPA: 3.83 *05/2020*