Shubham Nimbalkar

Sunnyvale, California 94086 | 949-539-6649

shubhamnimbalkar.com | srnimbal@uci.edu | Github | Leetcode

Technical Skills

Public cloud/on-prem, Docker, Kubernetes, Git, JIRA | Sagemaker, Lambda, Bedrock | Spark on K8s, Hudi | Tensorflow, Keras, Spark | Airflow, Kafka, Data pipeline management | Clickhouse, Mysql, Elastic Kibana, Superset, Splunk |

Professional Experience

Salesforce | *Senior AI Infra Engineer (SMTS)*   **Sep 24 - Present**
- Building multi-substrate cloud infrastructure for **Gen-AI** research and products

Tesla, Inc. | *Software Engineer*   **Feb 23 - Aug 24**
- Building **spark-on-kubernetes** on-prem infrastructure to horizontally scale Cell Manufacturing data jobs by 10x
- Maintain real-time analytics for 3 factories on the in-house dashboard that provides yields at different granularities
- Leveraging best practices with **Kafka**, **Spark**, **Hudi** and **Clickhouse** to optimise for latency and compute resources
- Designed data pipeline for genealogy mapping from a cell to raw materials across different shops in manufacturing
- Built long running streaming jobs with E2E tests and easy deployment that can consume millions of messages per hour

Apple, Inc. | *Software Engineering Intern*   **June 22 - Sep 22**
- Early adoption of internal multi-device frameworks to automate and migrate jobs in current test pipeline
- ShieldSense - A dashboard for teams to visualise xcode project coverage metrics to improve software quality

Rakuten | *Software Engineer*   **Sep 19 - July 21**
- Developed end-to-end Campaign Intelligence Platform to improve ad-based targeting and boost campaign sales by 5%
- Designed **data pipeline** and cloud infra using **spark**, **elastic** and **SQL** to fetch and co-relate 20 GB of data daily
- Built a frontend with **React** for marketing team with scalable performance using **Kubernetes** and **Docker**
- Analysed empirical data and competitor insights to expand user base and provide targeted incentives
- Created LSTM based **sales predictor** to visualise sales 2 weeks into the future with more than 70% accuracy

Rakuten R&D | *Research Trainee*   **Jan 19 - Aug 19**
- Face Liveness Detection for Rakuten Pay - A **face authentication projec**t to make secure payments by thwarting 2d and 3d masked attacks. Collected face depth data of 300 people by IR sensor using python to create LFW dataset
- Completed a research project under guidance of Research Scientist, Nithish Divakar to describe an image with tags ranked by visual importance. Designed **web application** to upload images and inference the **AI model**.
- Created an automated video metadata extraction for OTT platforms to reduce human annotation effort by 90%.
- Implemented a pipeline to process video and output good quality, distinct and scene classified posters of movies

Rakuten | *Intern*   **Sept 18 - Dec 18**
- Built Item bundle **recommendation engine** to boost cross-item selling & average basket size using TF-IDF
- Analysed merchant's selling history, tokenized japanese text and created a similarity scoring matrix for catalogue
- Developed a hierarchical server health **monitoring and alerting system using node-link graph**
- Reduced the server downtime and empowered engineers to take action before any server crash (Similar to APM sys)

Projects and Recognition

In-house Search Engine | **2022 Capstone Project**
- Developed a search engine using TF-IDF for search and Trie for autocomplete. Hosted on internal Linux env.
- Developed content summarization for easy navigation through transcribed videos and article search results

Plant Disease Detection Using CNN and GANs | Python, Image processing, App development
- Co-Authored and presented research paper at ICIRD 2018, Thailand (IEEE and Scopus)
- Winning project in ML/AI domain at INC event, PICT and Runners up at Quark event. BITS Pilani

Typist DNA | Python, Java, Machine Learning - An authentication system to verify password by analysing typing patterns with clustering techniques using keystroke metadata as data points

NDCG Metric Calculator | Python - Created a script to to evaluate NLG models accuracy using NDCG formula | Github

Face Crop - Extract frontal-face frames from a photo or video using haar cascade filter | Github

Education

University of California, Irvine
Master of Computer Science   **September 2021 - December 2022**
Pune Institute of Computer Technology
BE in Computer Science   **August 2014 - August 2018**