NIHAR JOSHI

Chicago, IL • +1 (312) 647-3243 • nsj0596@gmail.com • github.com/niharjoshi • linkedin.com/in/nihar-s-joshi

EDUCATION

Master of Science in Computer Science | **University of Illinois at Chicago** | 3.67 GPA — Aug 2021 – May 2023
Bachelor of Science in Information Technology | **National Institute of Technology, Karnataka** — Aug 2014 – May 2018

EXPERIENCE

Software Engineer II | **EagleView Technologies** | Bengaluru, India — Aug 2019 – Jun 2021

- Designed the EagleView Assess Machine Learning pipeline for real estate property inspection to serve more than 2 million customers (and clients such as Tesla Solar) with 98% customer satisfaction.
- Reduced real estate property inspection time from 3 days to 6 hours per property via automation using a stack of Python, Go, Kubernetes, AWS and Kafka, saving the company more than $70,000 annually in infrastructure costs.
- Converted legacy Image Processing APIs into robust and scalable gRPC Go services capable of handling 100 concurrent streams of requests per second while honoring a Service-Level Agreement of 99.999% uptime (5-nines Availability).
- Designed a Multi-threaded image tiling API in Python OpenCV to reduce the image retrieval time by 55% for the internal GraphQL API-based database housing more than 1 billion high-resolution property images.
- Assisted the Data Science team design a seamless CI/CD process to host various Anomaly Detection models orchestrated via Netflix Conductor to support 1000+ model predictions per second.
- Earned the Above & Beyond Performance award from the CEO for cross-team leadership, communication ability, presentation skills and driving EagleView Assess to production within 6 months.

Software Developer | **Graphene AI** | Bengaluru, India — Jun 2018 – Jun 2019

- Led the CI/CD and Inter-process Communication (IPC) development for the Mavis AI engine to automate market research for Procter & Gamble, Johnson & Johnson, Sanofi, Abbott and other top healthcare specific companies.
- Designed a distributed text processing module using Python, Apache Spark and Apache Kafka to automate data extraction, stopword removal, stemming, lemmatization and tokenization of more than 10 million lines of text per day.
- Developed a trio of NLP models in Python to perform Part-of-Speech (POS) Extraction, Named Entity Recognition (NER) and Aspect-based Sentiment Analysis (ABSA) of pharmaceutical data with 94% accuracy.
- Improved the QA process by refactoring underperforming IPC modules and increasing the integration test coverage of internal pipelines from 45% to 87%.
- Chosen as the Lead SWE to collaborate with Procter & Gamble by virtue of being the highest performing member of the AI Team.

PROJECTS

Contextual Object Relationship Identification for Anomaly Detection & Image Enhancement — May 2018

- An application that performs real-time Scene Recognition and Anomaly Detection using Faster-RCNN based Deep Learning to detect emergencies in live video feeds (example - fire in a house, accident on a road) with an accuracy of over 94%.
- Awarded the Best Student Project Award for 2018 by AMD for our contribution to the development of Smart Cities.
- Submitted the project paper to the 2019 Conference on Computer Vision and Pattern Recognition.

Log File Monitoring & Alert System — Dec 2021

- An Apache Spark based log file processor written in Scala that sends automated alerts to stakeholders based on log severity and frequency for cloud monitoring and troubleshooting.
- Consists of an Akka actor system deployed on an Elastic Kubernetes Service cluster with ElastiCache (Redis) database for persistent storage.
- Uses Amazon's Managed Streaming for Apache Kafka to perform inter-process communication and AWS SNS to send email notifications.

Text Data Classification Using MapReduce — Sept 2021

- A modular Parallel Processing application written in Scala that uses MapReduce to rapidly classify terabyte-level Big Data based on predefined metrics.
- Also consists of a Python Lambda function exposed via an AWS gRPC API Gateway endpoint to allow quick retrieval of classified logs in O(log n) time over O(1) space using Binary Search.

SKILLS

Proficient: Python, Bash, AWS, Docker, Kubernetes, Apache Kafka, Jenkins, Argo CD, Redis, PostgreSQL, Flask, CI/CD, REST APIs, Linux/UNIX, Shell Scripting, Git, Unit/Integration Testing, DSA
Exposure: Scala, Go, Apache Hadoop, Apache Spark, gRPC, OpenCV, GraphQL, Google Cloud Platform, Microsoft Azure