OMKAR KAKADE

omkar.kakade@gmail.com
o-kakade.github.io
5857555688
15 Quarry Ln, Malden, MA, 02148
omkarkakade
o-kakade

Skills

LANGUAGES
Python
Java
JavaScript
TypeScript

TOOLS
Git
Bash
Docker
Nginx

AI/ML
PyTorch
Pandas
Numpy
Scipy
Matplotlib
OpenCV
spaCy
scikit-learn
Tensorflow

DATABASES
MySQL
Blazegraph
AWS Neptune
SQL
SPARQL

CLOUD
Amazon Web Services - AWS
Google Cloud Platform - GCP
Microsoft Azure

WEB
Flask
Spring
Angular
Nest.js
HTML
CSS

Certifications

AWS Certified Solutions Architect - Associate — 2019 - 2022

AWS Certified Developer - Associate — 2020 - 2023

AWS Certified SysOps Administrator - Associate — 2020 - 2023

Education

Rochester Institute of Technology — Aug. 2017 - Dec. 2020
Master of Science Computer Science
GPA : 3.64

University of Pune — Aug. 2012 - May 2016
Bachelor of Engineering Computer Engineering

Experience

Software Engineer — Somerville, MA
Motorola Solutions — Nov. 2021 - Current

- Full Stack Software Engineer with CTO Systems and Solutions group
- **CAPE Drone Telepresence** - Implemented microservice to cleanup records from firestore that were prevent users from seeing new live flights. (*Python*, *Typescript*, *Angular*, *Django*, *PostgreSQL*, *Firestore*, *GCP*)
- Build full feature for Livestream Management to add/remove users from livestream. Implement APIs, Frontend components and Microservices to sync state across UIs with firebase.

Software Engineer - AI/ML — Boston, MA
Bola AI — Feb. 2021 - Oct. 2021

- **Voice Enabled Dental EHR System** - Designed and deployed cloud based infrastructure to scale Bola backend and support 3x growth of users. (*Azure*, *Docker*, *NGINX*, *App Gateway*, *Prometheus*, *Grafana*)
- Built CI/CD pipelines with configuration scripts on VMs to enable reliable deployment and release cycles, boosting engineering team productivity with sub 10 minutes deployment. (*Github Actions*)
- Developed a custom speech model in collaboration with Deepgram and integrated it with the backend to decrease response times - 2x faster, scale with 58% more cost efficiency and increase intent accuracy by 10% for users improving overall experience. (*Deepgram ASR*, *Typescript*, *Python*)
- Built tool to automate KPI workflows carried out by business team saving atleast 30+ hours per week on data triaging and analysis. (*Python*, *Azure SDK*, *Pandas*)
- Implemented feature toggling in the backend to conduct beta testing with live traffic, increasing reliability for new feature releases. (*Typescript*, *Nest.js*, *WebSockets*, *Angular*, *Electron*, *Sentry*)

Machine Learning Engineer — Orlando, FL
Siemens — Jan. 2020 - May 2020

- **Internal Predictive Analytics Platform** - Configured *RStudio Server Pro* with *NGINX* reverse proxy enabling a cloud based model development space.
- Built training and inference *Docker* images for custom R based time series forecasting use cases to enable scalable model training and model deployment.
- Developed python scripts using *Sagemaker SDK* for deploying models using batch transform to serve predictions on-demand enabling cost savings relative to deployment using always-on model endpoints.

Software Developer — Orlando, FL
Siemens — May 2019 - Dec. 2019

- **Search Application for Knowledge Graph** - Developed RESTful API for a keyword recognition based serverless search application to provide answers to user's questions from a *RDF graph database* in AWS. (*Python*, *Java*, *spaCy*, *Apache Jena*, *Flask*, *Springboot*, *AWS Lambda*, *API Gateway*, *Neptune*, *S3*, *Angular*)
- Designed and developed all modules with OOP from scratch such as Orchestrator, Keyword Recognizer, Indexer, Query Builder, Query Executor.
- **Database Migration** - Migrated production dump of graph database from Blazegraph to AWS Neptune to mitigate security concerns and move to a cloud based RDF triple store.
- **Cloud Resource Management** - Implemented API for a resource scheduler tool with features such as auto start-shutdown for a total cost saving of 87% relative to always-on EC2.

Projects

Performance and Deployment of Deep Neural Net on Edge Devices — 2020

- Deployed InceptionNet and MobileNet on Raspberry Pi 4 while measuring performance on metrics like accuracy, file size, CPU and Memory usage, latency. (*Python*, *TF-Lite*)
- Improved performance across all metrics while maintaining accuracy using techniques like quantization, weight clustering and weight pruning.

Multi-core, Cluster, GPU and Map-Reduce Projects - https://bit.ly/2IJBf8B — 2018

- Developed parallel computing programs to solve large mathematical problems demonstrating strong scaling and weak scaling. (*Parallel Java 2*, *Java*, *C*)

Transfer Learning using VGG 16 and LeNet 5 — 2018

- Adapted pre-trained models to work with Fashion MNIST dataset and improved accuracy by fine tuning the networks to yield accuracy of 91% with VGG-16 and 86% with LeNet5 pre-trained on MNIST. (*PyTorch*, *Matplotlib*)