Email: raman.shinde15@gmail.com https://github.com/Raman-Raje Mobile: +919595161238

Career Summary

- Senior Lead Engineer with expertise in Machine Learning compilers, with a focus on TVM, enhancing network performance by optimizing code generation and runtime components.
- Expertise in developing Deep Learning models using architectures such as LSTM/RNN, CNN, Transformers, Autoencoders and LLM's.
- Experience with all major ML algorithms, including KNN, K-means, Naive Bayes, Logistic Regression, SVM, Decision Trees, Random Forest, and GBDT.
- Proficient in leveraging cloud platforms (AWS, GCP) and containerization/orchestration tools like Docker and Kubernetes for scalable ML deployments.

EXPERIENCE

Qualcomm

Banglore, India

Mar. 2024 - Present

Senior Lead Engineer [AI Research]

- * Implemented optimized data handling for Adreno GPUs in TVM by adding texture support for OpenCL and Vulkan backends, resulting in a 50% improvement in network inference times.
- * Enhanced TVM capabilities by adding support for new layers and operators to accommodate the latest Large Language Models (LLMs) and Diffusion networks on Adreno GPUs.
- * Incorporated memory planner to support and efficiently run large-scale LLMs within constrained memory limits.

• Imagination Technologies

Pune, India

Deep Learning Engineer

Sep. 2021 - Mar. 2024

- * Developed an SDK enabling neural network execution on Imagination's GPUs and Neural Network Accelerators (NNAs) using TVM, facilitating efficient model deployment on edge devices.
- * Implemented LSTM/RNN support in Imagination's Neural Compute SDK (NCSDK) for models built with PyTorch, TensorFlow, and ONNX, ensuring broader compatibility and accelerated performance.
- * Created and optimized graph transforms for diverse operations and operators within RelayIR (TVM), enhancing computational efficiency.
- * Contributed to the development of quantization tools (static, dynamic, and QAT) for multiple frameworks, driving improvements in model compression and inference speed.

 Xpanxion Pune, India

Data Scientist

Jan. 2020 - Sep. 2021

- * Developed AI/ML solutions to extract information from medical documents, including document classification and candidate extraction (e.g., benefits, rates, and drug details) using libraries like Fonduer, Tesseract, and OpenCV.
- * Built reusable AI/ML components within the Innovation Team, including a content-based and collaborative recommendation system, NLP modules such as Named Entity Recognition (NER), QA systems, and sequence translation.
- * Led Computer Vision projects involving object localization and detection, image segmentation, and gesture recognition.

• Siemens R&D Pune, India

Product Development Engineer

Dec. 2018 - Dec 2019

- * Developed and enhanced an application for designing manufacturing sequences, with significant contributions to debugging and resolving issues in existing code.
- * Accelerated feature deployment by implementing new use-cases and Proof-of-Concepts (POCs), enabling faster shipping of key features.

• TCS Pune, India Dec 2015 - Nov 2018

Software Developer

* Developed monitoring and control applications for NCRA, including real-time issue monitoring, debugging, and GUI modifications based on client requirements.

- * Provided support for financial applications such as CRD and SRD for Morgan Stanley, ensuring reliable performance and issue resolution.
- * Built applications for various clients using Python and C++, contributing to diverse projects and delivering tailored software solutions.

TECHNICAL SKILLS

• Languages: Python, C++

• Database: MySql, MongoDB

• Data Analysis: Pandas, Numpy, Matplotlib, Seaborn, openCV

• ML/DL Toolkit: Keras, scikit-learn, tensorflow, pytorch, TVM

CERTIFICATIONS/INTERNSHIP

• Applied Machine Learning course at Applied AI. (Jan 2018 to May 2019)

- Completed Standford Statistical Learning (Self-Paced) course.
- Completed Deep Learning Specialization course from Coursera
- Internship at IARE, Aurangabad on Industrial automation. (May 2014 Jun 2014)

EDUCATION

- B.Tech in Electronics and Telecommunication from SGGSIE&T, Nanded with CGPA 7.7 (2011 2015)
- Class Xll (HSC), form Maharashtra State Board of Education with 83.33% (2009 2011)
- Class X (SSC), form Maharashtra State Board of Education with 90.92% (2008 2009)