Yash Akhauri

Blog | GitHub | LinkedIn | Gmail | +91 78915 12802

EDUCATION

BITS Pilani
B.E. Electronics and Instrumentation
Expected May 2020

Ryan Global School
IGCSE 2014 | Kharghar, India
Percentage 90.1

SKILLS

Programming languages:
- Python
- C++
- Java
- Mathematica

Others:
- CUDA
- libGDX
- Linux
- Tensorflow
- OpenMP
- Android studio
- Excel
- PyTorch

EXPERIENCE

WOLFRAM | Undergraduate Researcher
June 2018 - July 2018 | Waltham, Massachusetts

- Conducting research on distributed mesh AI computing by introducing Hadamard Binary Neural Networks.
- Developed optimized C OpenMP kernels for GEMM and convolution using the Hadamard Binarization methodology.
- Implemented MLPs for full precision, binary precision, and hadamard binary neural networks from scratch in the Wolfram Language.

LARSEN & TOUBRO | Intern
December 2016 | India

- Scraped and managed stock market data from BSE.
- Developed and tested several machine learning techniques to analyse stocks and developed investment strategies.

DEPARTMENT OF SPONSORSHIP | Core member
August 2016 – Present | India

- Department Of Sponsorship and Marketing looks after the budget of the technical fest of BITS Pilani, APOGEE and cultural fest of BITS Pilani, OASIS.
- Have made calls for sponsorship and media relations, handled sponsors on campus and brainstormed over branding avenues and marketing options.

GRANTS AND TALKS

HadaNets: Flexible Quantization Strategies for Neural Networks
Paper under review at a conference.

Wolfram Technology Conference – Speaker Champaign, IL | Oct. 2018
Delieverd a talk on my research on Hadamard Neural Networks.

Intel AI Meetup – Speaker [pptx] Delhi, IN | Sept. 2018
Spoke about my research on scaling AI using Intel techonologies. This event was jointly organized by Intel Software and Analytics Vidhya

Intel AI Academy Success Story [Link]
Intel published a cover story of my research done in the field of Quantized Neural Networks – Convoluted Neural Network Optimization

Intel AI DevCon – Poster May & Aug 2018
Presented a poster on high performance binary neural network kernels for Intel Xeon Phi at Intel AI Student Summit, San Francisco in May and Intel AI DevCon, India in August.

Intel Nervana Early Innovators Grant $5000
Received research grant to develop Binary Precision Neural Networks and Real time Artistic Style Transfer.

Wolfram Research Grant $2400
Received aid to attend the Wolfram Summer School and develop Hadamard Binary Neural Networks.

KVPY Scholar
Selected as a KVPY scholar by the Department of Science and Technology, Government of India.

PROJECTS

Hadamard Binary neural networks — C++, CUDA, OpenMP, Mathematica

Developing a neural network architecture that is meant to provide low latency, low powered solution for distributed mesh AI computing and offline inference. [Performance and accuracy analysis]

- Increased network inference speed 7x with only a 2% decrease in accuracy.
- Studied the High-Dimensional geometry of the new binarization scheme.
- Developed optimized GEMM and convolutional kernels for HBNN MLPs using OpenMP.

Whitepaper - Improving distributed mesh computing with Hadamard Binary Neural Networks.

[Link]

xGEMM & xCONV — C++, CUDA, OpenMP

Coded efficient 3D convolutional and GEMM kernels for XNOR (bit quantized) networks using CUDA C programming and OpenMP. Optimized kernels are for Intel processors and Nvidia GPUs. Invited to present a poster at Intel AI DevCon and Intel AI Student Ambassador Summit, San Francisco. The codes can be found here: [OpenMP kernel] | [CUDA kernel].

GEMM: From Pure C to SSE Optimized Micro Kernels — C++

Followed this tutorial to build an implementation of GEMM that achieves performance close to that of the BLIS kernels.

Real time artistic style transfer — Python, Tensorflow, OpenCV

Developed an activation based Adaptive instance normalization technique. Worked on a pruned generator network with Instance normalization to give stylization speeds of 25 FPS at VR resolution. Integrated pruned network with Android IP Webcam for live stylization of camera feed. The technical article can be found [here.] The code can be found [here.]

GravDash — Java, LibGDX, Android Studio

Developed an android game using Java and libGDX as the framework. The game can be found [here.]

Blog — Python, Java, C++, Tensorflow, OpenMP

Maintaining a [blog] covering various AI related topics with over 10000 hits.