Varun Desai

vdesai@stanford.edu linkedin.com/in/vdesais github.com/varundesai10

Education

Stanford University **September 2023 – Present**

Master of Science in Electrical Engineering (Hardware and Software Systems); ***GPA: 4.3/4.0*** *Stanford, CA*

- Coursework: Operating Systems, Embedded Systems Workshop, Parallel Computing, Deep Generative Models

Indian Institute of Technology, Delhi **July 2018 – April 2022**

Bachelor of Technology in Electrical Engineering; ***GPA: 9.625/10*** *(1st among 120 students)* *New Delhi, India*

- Coursework: Computer Architecture, Intro to Machine Learning, Data Structures and Algorithms, Synthesis of Digital Systems, Planning and Estimation for Autonomous Systems, Convex Optimization, Probability and Stochastic Processes.

Achievements

Institute Silver Medal, IIT Delhi: For attaining the highest CGPA among distinguished EE graduates. **2022**

KC Mahindra Scholarship : awarded for exceptional academic excellence, to support graduate education. **2023**

IIT Delhi Merit Award : 5 times in 8 semesters for consistently outstanding academic standing **2018-2022**

JEE Advanced, 2018 : Attained All India Rank of **301** among 150,000 candidates, JEE Mains AIR 2048 **2018**

Work Experience

Graviton Research Capital, LLP **June 2022 – June 2023**

Quantitative Trader/Researcher *Gurugram, India*

- Spearheaded end-to-end deployment, management and refinement of latency-sensitive trading strategy for index options
- Set up automated post-trade analysis of big datasets of recorded market data using C++, python revealing performance gaps and identifying avenues for enhancing market-making efficiency while mitigating multifaceted risk.

APT Portfolio **May 2021 – July 2021**

Hardware Verification Intern *Bengaluru, India*

- Designed, tested architecture for verifying RTL model on an FPGA to speed up verification process more than 100x
- Developed digital circuit using Verilog to (de)serialise data to and from compressed format for efficient transmission
- Used Riviera-PRO and Cocotb to debug and test design, Xilinx Vivado for synthesis and ensuring timing requirements.

Greenleap Robotics **May 2020 – July 2020**

Computer Vision Intern *New Delhi, India*

- Implemented ML based Computer Vision techniques to automate fault diagnosis of Solar Panels via Drone Imagery

Projects & Publications

Working on the Pintos Operating System | *CS212, Stanford* **January 2024**

- Implemented kernel threads equipped with a multi level feedback queue scheduler in the Pintos operating system
- Argument passing on stack implemented for user programs, along with various system calls such as exec, wait, open, read

Flash Attention for NanoGPT Implementation | *CS149, Stanford* **December 2023**

- Focused on enhancing the performance of the attention layer of the model in C++; Employed loop blocking, loop fusion to optimize arithmetic intensity and reduce memory footprint. Leveraged OpenMP and ISPC for further speedup.
- Flash Attention implemented to break softmax into blocks, allowing for fusion with blocked matrix multiply.

Targeted Machine Unlearning in Diffusion Models | *CS236 Course Project, Stanford* **December 2023**

- Pioneered the integration of certified robustness with adversarial attack mechanisms to rigorously assess efficacy of targeted unlearning in SoTA diffusion models; Developed and applied novel techniques using Stochastic Gradient Ascent and Influence Functions for selectively erasing data points from complex neural network architectures.

On-chip Learning of a Convolutional Neural Network | *Neuromorphic Comp. & Engg. 2022 vol 2* **December 2021**

- Designed crossbar array architecture for novel neuromorphic synaptic device to facilitate on-chip learning for a deep convolutional neutral network; Analysis of robustness against noise and other non idealities also performed.
- Developed novel training algorithm to deal with limited bit precision available on the synaptic devices.

Planning and Estimation for Autonomous Systems | *COL864, IIT Delhi* **March 2021**

- Implemented value iteration for MDPs, RL algorithms(SARSA,Q-learning); Viterbi Algorithm for HMMs
- Devised algorithm for estimating position of multiple agents with no data association as extension to Kalman Filtering

Languages & Technologies

C++, C, CUDA, OpenMP, Python (pandas, NumPy, PyTorch, TensorFlow), R, Java, Assembly (ARM), Verilog, VHDL, MATLAB, Git, LaTeX, Linux, bash, Arduino