

# Prashant Sohani

< [psohani@nvidia.com](mailto:psohani@nvidia.com) >

Software engineer with 14 years of experience in applied R&D of algorithms and systems

## Work experience

---

### NVIDIA Corporation

Munich, Germany

*Senior AI Developer Technology Engineer*

April 2024 – present

CUDA optimization of datacenter-scale training and inference models, with a focus on MLPerf and Genomics

- OpenFold2 inference optimization
  - Using TensorRT and GPU pipelining to accelerate protein structure prediction by [more than 100x](#)
- AlphaFold Complex Database generation
  - Collaboration with multiple organizations to expand the database with [high-confidence multimers](#)

### NVIDIA Corporation

Pune, India

*Senior Systems Software Engineer*

October 2018 – March 2024

*Systems Software Engineer*

July 2012 – September 2018

Applied research for algorithms and ML models in video processing, network streaming, and 3D graphics

- Reinforcement Learning for QoS policies in Geforce Now
  - Developed [patented algorithms](#) for bitrate control to [improve or replace](#) conventional QoS methods
- Deep Learning algorithms for video quality improvement
  - Achieved real-time HD upscaling for game streaming, by training and pruning GAN models
- Weighted Prediction for NVENC video encoder
  - Delivered BD-PSNR gains of over 1dB, driving the feature cycle from pre-silicon to driver integration
- Camera image noise reduction
  - Developed [patented algorithms](#) for low complexity filtering, for denoising low-light mobile captures

Other activities at NVIDIA: NTech internal conferences, site-level ML training workshops, university campus hiring

### Oneirix Engineering Labs

Pune, India

*Software Engineering Intern*

May – July 2009, May – July 2010

Leveraged concepts in audio processing and queuing theory, for echo cancellation and network simulation systems

Developed Linux user-space applications as well as kernel modules, using ALSA and netfilter frameworks

## Education

---

### Indian Institute of Technology Bombay

Mumbai, India

*B.Tech. in Electrical Engineering*

July 2008 – April 2012

*with minors in Industrial Design*

Grade: 8.0 / 10.0

Undergraduate Research Award (URA 01) for B.Tech. thesis on Parameter Estimation of Chirp Signals

Bronze medalist at the **International Mathematical Olympiad 2008**, held in Madrid, Spain

## Technical skills

---

Programming languages and frameworks:

Python, C/C++, CUDA, TensorRT, ns-3, OpenCV

Python for ML, statistics, and graphing:

PyTorch, TensorFlow, NumPy, Pandas, Matplotlib

3D rendering APIs, game engines and tools:

OpenGL, HLSL, Unreal, Unity, Godot, Blender

Linux internals and subsystems:

ALSA, networking and IPC, kernel module

Full-stack web technologies:

Apache, MySQL, PHP, NodeJS, MongoDB, React

## Other activities

---

- Over 2000 hours of training experience at various Mathematical Olympiad camps since 2008
- Regional Coordinator of Mathematical Olympiads for Maharashtra and Goa from 2016 to 2022
- Playing chess and music since childhood, self-trained to play the keyboard and guitar