

Prateek Ghosh

San Francisco, CA – USA

+1 (716) 994-2633 | prateek.prateekghosh29@gmail.com | www.linkedin.com/in/prateek-ghosh-ub/
github.com/Prateek-2106 | prateek-2106.github.io

EXPERIENCE

School of Engineering and Applied Sciences, University at Buffalo

Buffalo, NY, USA

Research Intern

Oct 2025 – Feb 2026

- Achieved a 4.44 Mean Opinion Score (MOS) by curating a DiffWave pipeline for speech generation using mel-spectrogram conditioning and reverse diffusion training.
- Enhanced speech signal enhancement models by experimenting with LSTM and GRU architectures, resulting in surpassing WaveNet performance by 43%.
- Delivered 100% data completion by implementing a diffusion process with GNN and FGTI for vehicle speed imputation using PyTorch and Hugging Face.
- Processed and analyzed 100,000+ vehicle speed data samples, mitigating outliers and incompleteness due to weather conditions.

Deloitte US Offices of India Pvt Ltd

Hyderabad, India

Software Developer – SAP MDG

Jan 2023 – Jul 2025

- Engineered a PII data-scrambling compliance automation platform using Python and SAP ABAP, automating manual data-masking workflows and improving enterprise data-privacy compliance across CRM systems by 90%.
- Refactored legacy cloud-based systems as part of the Custom Code Decommission initiative, decommissioning over 35,000 redundant objects and improving system response time by a substantial 45
- Earned Deloitte's Applause Award due to technical excellence and cross-functional collaboration on enterprise systems

Software Developer – SAP ABAP

Jul 2021 – Dec 2022

- Pioneered the implementation of 110+ production-grade model enhancements via Agile and CI/CD pipelines, while adhering to cross-team SLAs to 87% through test-driven feature development.
- Spearheaded the optimization of SQL queries, improving data retrieval speed by 10% while resolving 70+ critical bugs.

Central Tool Room & Training Centre (Govt. of India)

Bhubaneswar, India

ML & AI Intern

Nov 2019 – Dec 2019

- Trained a real-time human-figure detection model using TensorFlow CNNs; marked 92% accuracy, reduced image-processing latency by 22%, and lowered false-positive rate by 8% through model tuning.
- Benchmarked classical ML models (KNN, SVM, Naïve Bayes) against deep-learning baselines; achieved 10%+ accuracy improvement through model tuning and dataset preprocessing pipelines using PyTorch.

Tata Motors South Africa Pvt Ltd

Pretoria, South Africa

Software Developer Intern

May 2019 – Jun 2019

- Developed an end-to-end automation system integrating Excel VBA, SQL, and live scanner data, enhancing production-line traceability for vehicular manufacturing, increasing production from 11 to 13 units daily.
- Automated the inventory-tally workflow, cutting cycle time by 90% and query latency by 98% across 9 vehicle models.
- Replaced external vendor software with an in-house data management platform, saving ZAR 15,000 annually and improving scalability, maintenance, and operational control.

ACADEMIC RESEARCH & PROJECTS

Ozone Concentration Time-Series Forecasting

Feb 2026

- Designed a custom MLP Neural Network trained for time-series forecasting of Ozone Concentrations at time $(t + k)$ hours using a USA gov hourly meteorological dataset with PyTorch.
- Implemented look-back and lag features to reveal long-term patterns across 78 sites after data cleaning, imputation, and EDA; observed performance comparable to LSTM and GRU baselines.

Simulation of Autonomous Driving

Dec 2019

- Delivered 92% accuracy with a custom CNN in Keras for steering prediction using 10,000+ manually collected images; sustained 95%+ lane-keeping accuracy and decreased steering error variance by 30%.

EDUCATION

University at Buffalo, SUNY

M.S., Computer Science (AI/ML Track)

Aug 2025 – Dec 2026

- Coursework:** Algorithm Analysis and Design, Data Intensive Computing, Computer Security

KIIT University

B.Tech, Electronics & Computer Science

Jul 2017 – Aug 2021

- Coursework:** Data Structures and Algorithms, Object-Oriented Programming, Operating Systems

SKILLS

Languages: Python, C, C++, MATLAB, JavaScript, Node.js, HTML, CSS

Databases: MySQL, PostgreSQL, MongoDB, Snowflake

Frameworks & Libraries: TensorFlow, PyTorch, Keras, Scikit-learn, Hugging Face, Transformers

Core Competencies: Diffusion Models, Deep Learning, NLP, CNNs, Computer Vision, Time Series, LLMs, Mathematics, Statistics, Data Science, Data Visualization

Cloud & DevOps: AWS, GCP, Docker, Git, Bash, Spark, Hadoop, REST APIs, CI/CD, Linux

Enterprise Tech: SAP ABAP, SAP MDG, Power BI, ETL Processes, Advanced Analytics