Prateek Ghosh

Buffalo, NY - USA

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

github.com/Prateek-2106

Summary

Machine Learning focused MSCS candidate with 3+ years of experience in Python and ML/AI. Proven expertise in data cleaning, transformation, and feature engineering, including building a PII-compliance automation platform to enhance data privacy. Developed and tuned TensorFlow CNNs and classical ML models achieving 92% accuracy, and automated data workflows using SQL and Python. Skilled in pandas, NumPy, scikit-learn, and collaborating with cross-functional teams to deploy AI-driven features.

Experience

School of Engineering and Applied Sciences, University at Buffalo **Oct 2025 - Present**

Research Intern

- Reported 4.44 Mean Opinion Score (MOS) in Speech Audio Synthesis (DiffWave Implementation): Curating the DiffWave pipeline for speech generation using the pre-trained 22.05 kHz model with mel-spectrogram conditioning; inject Gaussian noise via a Markov chain forward process and train reverse diffusion to recover clean audio.
- Vehicle Speed Imputation (Diffusion + GNN/FGTI): Studying and applying the above diffusion process to Graph Neural Networks (GNN) and FGTI for imputing missing values in Buffalo roadside speed-sensor time series; coverage windows: Oct-mid Nov 2022, first half of Dec 2022, Feb 2023; imputation targets: half of Nov 2022, second half of Dec 2022, Jan 2023 (full).

Deloitte US Offices of India Pvt Ltd **Jul 2021 - Jul 2025**

Software Engineer (SAP, Cloud, Data Systems)

- Engineered a PII data-scrambling and 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 in the Custom Code Decommission initiative; decommissioned 35K+ redundant objects, improved system response time by 45%; enabled a $350K project renewal through scalable backend redesign.
- Delivered 110+ production-grade enhancements under Agile and CI/CD pipelines in ServiceOne; strengthened cross-team SLA adherence to 87% through test-driven, maintainable feature development.
- Resolved 70+ ABAP defects using SQL tuning, debugging, and OOP principles, boosting query processing efficiency by 6%.
- Recognized with Deloitte's Applause Award for technical excellence and Cross-functional Collaboration on enterprise systems.

Central Tool Room & Training Centre (Govt. of India) **Nov 2019 - Dec 2019**

ML & AI Intern

- Trained a real-time human-figure detection model using TensorFlow CNNs; achieved 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 efficient model tuning and dataset preprocessing pipelines using AI/ML frameworks (PyTorch).

Tata Motors South Africa Pvt Ltd **May 2019 - Jun 2019**

Software Developer Intern

- Developed an end-to-end automation system integrating Excel VBA, SQL, and live scanner data, enhancing production-line traceability and shop-floor preprocessing for vehicular manufacturing, increasing production from 11 units 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

Resume Classification

- Developed a role-matching model using Python (NLTK, Scikit-learn) to classify resumes by job family. Reached 83% model performance in offline evaluation.

Simulation of Autonomous Driving

- Collected and preprocessed 10,000+ images with augmentation to expand dataset diversity 3x.
- Delivered 92% validation accuracy with a custom CNN in TensorFlow for steering prediction with a custom CNN. Sustained 95%+ lane-keeping accuracy and decreased steering error variance by 30%.

Education

University at Buffalo, SUNY **Aug 2025 - Dec 2026**

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

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

KIIT University **Jul 2017 - Aug 2021**

B.Tech, Electronics & Computer Science

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

Skills

- **Programming Languages**: Python, C, C++, MATLAB, JavaScript, Node.js, HTML, CSS
- **Core Competencies**: Diffusion models, Deep Learning, NLP, CNNs, Computer Vision, Time Series, LLMs, PCA, Mathematics, Statistics, Data Science, Data Visualization
- **Frameworks and Libraries**: TensorFlow, PyTorch, Keras, Scikit-learn, Hugging Face, Transformers
- **Databases**: MySQL, PostgreSQL, MongoDB, Snowflake
- **Cloud & DevOps**: AWS, GCP, Docker, Git, Bash, Spark, Hadoop, REST APIs, CI/CD pipelines, Linux
- **Enterprise Tech**: SAP ABAP, SAP MDG, Advanced Analytics, Cross-functional collaboration, ETL Processes, Power BI