Utsav Shah

Washington, D.C. | (202)-569 0857 | utsavshah944@gmail.com | linkedin.com/in/shah-u/ | github.com/shah-utsav/ | Portfolio

EDUCATION

Howard University (CGPA: 4.0/4.0)   Washington, D.C.
Bachelor of Science in Computer Science (Minor: Mathematics)   *Aug 2023 – May 2027*
- Relevant Coursework: Computer Science I, Algorithms & Data Structures, Object-Oriented Programming (OOP), Linux Lab, Calculus I-III, Linear Algebra, Discrete Structures, Computer Organization, Applied Data Science, Introduction to Cybersecurity, Operating Systems, Physics I

EXPERIENCES

Mathematics Peer Tutor   Aug 2025 – Present
Howard University   *Washington, DC*
- Tutored 20+ undergraduate students (with a 4.7/5.0 rating) in College Algebra I–II, Calculus I, and Applied Calculus, strengthening their conceptual understanding, analytical reasoning, and problem-solving skills through weekly one-on-one and group sessions
- Designed customized practice problems, review materials, and targeted learning strategies in collaboration with faculty to identify and support at-risk students, leading to improved comprehension and engagement
- Enhanced overall course outcomes and exam readiness, with measurable improvements in student test performance and average scores across multiple test, quizzes and mid-term examinations

AI/ML Undergraduate Research Assistant   Aug 2025 – Present
Howard University   *Washington, DC*
- Designed and developed a dynamic, location-aware advertising system using Node2Vec graph embeddings and K-Nearest Neighbor (KNN) to recommend context-specific ads on moving taxi displays in real time along Washington, D.C. routes
- Implemented graph-based models capable of both learning and selective unlearning, enabling privacy-compliant removal of nodes or edges without full retraining
- Collaborated with Dr. Legand Burge at Howard University to build data pipelines, integrate GPS/business datasets, and prototype real-time inference for on-vehicle deployment

Undergraduate Summer Research Intern   Jun 2024 – Jul 2024
University of Maryland   *College Park, MD*
- Reduced lunar dust accumulation by 70% by engineering and testing an innovative cleaning system, building a custom GUI using wxPython to control cleaning cycles and monitor device performance in real time
- Enabled 10-minute cleaning cycles and reusability across 15+ production users, by developing Python automation scripts with NumPy and Pandas to analyze sensor data, generate real-time graphs, and optimize the gecko-inspired roller device parameters
- Developed a robotic-compatible cleaning solution for NASA's Artemis Mission Lunar Surface Technology Research (LuSTR) project, integrating simulation models, control logic, and a wxPython-based interface to demonstrate feasibility for automated lunar surface maintenance

Nanomaterials Physics Undergraduate Research Assistant   Feb 2024 – May 2024
Howard University   *Washington, DC*
- Conducted DFT simulations and modeling of piezoelectric quantum dots (ZnO & ZnS), analyzing atomic/molecular structures using Siesta and Quantum ESPRESSO
- Improved quantum dot stress-response models by 25% accuracy, refining 100+ parameters to optimize mechanical stability
- Developed research contributions on quantum dot usability for quantum processors through cross-functional physics team collaboration

PROJECTS

ForgeSentinel — AI Deployment and Security Research Project   (Link)
Python, FastAPI, PyTorch, Hugging Face Transformers & Accelerate, BitsAndBytes, CUDA, Redis, Docker, Custom SentinelLayer modules
- Conducted experimental research on LLM inference optimization, evaluating quantization, caching, and batch serving to reduce inference cost and latency
- Developed benchmarking pipelines and visualization tools to analyze throughput, memory footprint, and latency trade-offs across Large Language Model configurations
- Designed prototype AI-native cybersecurity modules (SentinelLayer) to detect prompt injections, adversarial inputs, and data anomalies in deployed models
- Investigated the intersection of MLOps, optimization, and secure AI deployment, proposing a scalable framework for low-cost, secure AI inference

Real-time Order Processing System with Sub-second Latency   (Link)
Laravel, Livewire, Alpine.js, ApexCharts.js, PHP, Trix Editor, Cloudflare CDN, Progressice Web Apps (PWA), Apache
- Delivered high-performance real-time order processing system with sub-second response times across 4 core modules (digital menus, order tracking, reservations, inventory) by engineering scalable backend architecture
- Accelerated order processing workflows by 60% reducing manual handling time from 5 to 2 minutes per order across 19 restaurant locations through real-time kitchen coordination system
- Achieved 100% system uptime for reservation management with 87% improvement in customer retention rates by designing fault-tolerant architecture with redundant servers
- Enhanced operational decision-making efficiency by 40% through comprehensive analytics dashboard with live data visualization using ApexCharts.js and automated sales insights pipelines

High-throughput Client Data Pipeline with Concurrent User Handling   (Link)
PHP, CodeIgniter 3, MySQL, jQuery, Bootstrap 3, DataTables, TinyMCE, PHPMailer, FullCalendar, Dropzone.js
- Engineered high-throughput client data processing pipeline with real-time lifecycle management serving 100+ concurrent small business users through scalable microservices architecture
- Accelerated sales pipeline responsiveness by 50% improving client follow-up rates from 40% to 90% through sub-second data retrieval and real-time event-driven notifications
- Reduced system onboarding latency by 35% decreasing user setup completion from 20 to 13 minutes per client through intuitive interfaces and automated task scheduling
- Enhanced forecasting system performance by 30% through improved sales prediction accuracy via live data aggregation pipelines and predictive modeling algorithms

TECHNICAL SKILLS

Programming Languages: Python, Java, C/C++, SQL
Web & Cloud Technologies: HTML, CSS, JavaScript, TypeScript, RESTful API, ReactJS, AngularJS, Node.js, Next.js, Amazon Web Services (AWS), Firebase, Docker, Kubernetes, Vercel, Netlify
Development & Design Tools: Git Version Control System, GitHub, Linux/Unix, Figma, Adobe Photoshop, Canva, Arduino, Raspberry Pi
Data Science & Specialized: PyTorch, Scikit-Learn, OpenCV, Jupyter Notebook, Ms-Excel, HuggingFace, Matplotlib, Seaborn, NI LabVIEW, AutoDesk Fusion 360