Sunny Shah

734-272-5306 | sunnysha@umich.edu | linkedin.com/in/sunnypshah | github.com/sunnypshah1 | sunnypshah1.github.io

Education

University of Michigan **Expected: May 2027**

Bachelor of Science in Computer Science and Pure Mathematics *Ann Arbor, MI*

Relevant Coursework: Data Structures and Algorithms, Computer Science Theory, Discrete Math, Machine Learning, Computer Organization, Probability Theory, Object-Oriented Design, Operating Systems, Networking

Experience

Data & Artificial Intelligence Engineering Intern **May 2025 – Present**

Protiviti *Chicago, IL*

- Architect a Streamlit vendor-spend platform that ingests ZIP/CSV uploads, validates schemas, and outputs clean model-ready datasets, driving real-time FP&A insights for executives across Fortune 500 client teams
- Build a robust data-validation engine that auto-coerces dtypes, imputes nulls, and blocks bad records, eliminating downstream model failures and accelerating onboarding for eight enterprise ledger data sources
- Engineer feature pipelines for Random-Forest classification, Prophet forecasting, and Isolation-Forest anomaly detection, producing fully numeric training matrices in under two seconds on commodity hardware for 100k+ rows
- Train and serialize production-grade models (500-tree RF, tuned Prophet, 300-tree IF) achieving 90% category accuracy and sub-10% 12-month forecast MAPE on validation data while enabling seamless CI/CD deployment
- Deliver interactive Streamlit dashboards that surface top-25 anomalies, 12-month spend forecasts, and live classification scores, empowering analysts to uncover savings and risk within seconds during weekly executive demos
- Optimize ingestion throughput 3× via PyArrow parsing and session-state caching while maintaining minimal memory footprint, ensuring smooth performance on enterprise-scale datasets in cloud-hosted VDI environments

Software Engineer Intern - Research **November 2023 – August 2024**

Wayne State University - College of Engineering *Detroit, MI*

- Developed a drone design in CAD software, improving lift-induced drag and air resistance by conducting iterative tests and simulations, adjusting wing sizes in NVIDIA Omniverse to accurately model and evaluate real-world performance
- Utilized Power BI and Python scripts to visualize test data and create dynamic dashboards for model comparison
- Authored research proposal, outlining project goals, methodology, and expected outcomes, secured funding over $1,000
- Organized and analyzed over 20 large datasets by implementing advanced NumPy techniques, including vectorized operations, multidimensional array manipulations, and statistical analyses, significantly reducing processing time

Projects

Multi-Class Image Classifier | *Python, PyTorch, Neural Networks, Logistic Regression, NumPy* **March 2025**

- Built a complete multi-class image classification pipeline in PyTorch to identify over 10 different dog breeds from real-world image datasets, achieving over 92% test accuracy after extensive tuning, experimentation, and evaluation
- Preprocessed and batched more than 10,000 labeled dog images by applying normalization, reshaping, and creating efficient, GPU-accelerated DataLoader classes to enable smooth model training and performance benchmarking
- Implemented and evaluated both a logistic regression baseline and a fully connected neural network; reduced training loss by 85% over 50 epochs through use of stochastic gradient descent and careful hyperparameter tuning
- Mitigated overfitting and improved generalization using dropout (p=0.5) and L2 regularization (λ=0.01), leading to an 11% boost in validation accuracy and significantly improved stability across multiple runs

Business Website Development – My Hydro Depot | *React, AWS, SQL, SEO* **July 2020 – July 2022**

- Designed and developed a responsive e-commerce website with a React frontend, featuring dynamic product displays, category filters, and an intuitive user interface tailored for retail customers
- Built a backend data layer using SQL to store and manage product data including inventory, pricing, and descriptions; implemented CRUD operations to enable seamless updates and ensure data consistency
- Deployed the site to AWS with distributed hosting and configured performance settings, improving uptime reliability and boosting PageSpeed Insights score by 23 points through caching and optimization techniques

Technical Skills

Languages: Python, C++, SQL, JavaScript, HTML/CSS, LaTeX
Frameworks: React, Bootstrap, jQuery, PyTorch, OpenCV, Agile, Jira
Developer Tools: Google Cloud Platform (GCP), Amazon Web Services (AWS), Git, VS Code, Jupyter Notebook, Linux
Libraries: pandas, NumPy, Matplotlib, PyTorch, SciPy, Scikit-learn
Concepts: Machine Learning, Volatility Arbitrage, Delta Hedging, Regression Analysis, Neural Networks, Optimization, Market Making, Synthetic Arbitrage, Cache Optimization