Sankalp Pandey

sankalpp@uark.edu | (479) 250-2959 | github.com/sanpdy | linkedin.com/in/sanpdy | sankalppandey.tech

Skills

Languages: Python, Java, C++, JavaScript/TypeScript, SQL, PHP, Bash
Frameworks and Libraries: TensorFlow, PyTorch, Scikit-learn, NumPy, Pandas, React, Next.js, SvelteKit
Tools and Platforms: AWS, Microsoft Azure, Docker, GitHub Actions, MySQL, PostgreSQL, Vercel
Expertise: Automated ML Pipelines, Data Structures & Algorithms, Distributed Systems, CI/CD
Interests: Quantum Materials, Computational Neuroscience, Computer Vision, HealthTech, Generative AI

Education

University of Arkansas Fayetteville, AR
B.S. in Computer Science and B.S. in Computer Engineering, GPA: 3.78/4.0 *Aug. 2022 – May 2026*

- **Minors:** Data Analytics, Mathematics
- **Activities:** Association for Computing Machinery (ACM), IEEE, Honors College Research Grant Scholar
- **Relevant Coursework:** Artificial Intelligence, Computer Vision, Data Mining, Algorithms, Software Engineering, Probability and Statistics, Linear Algebra

Professional Experience

Hidalga Technologies Springdale, AR
Software Engineering Intern *Aug. 2024 – Present*

- Developed a gradient boosting model with 87.6% accuracy on 30K+ cases to predict prior authorization outcomes, enabling smarter prioritization and reducing human reviews by $\approx 30\%$
- Piloted an automated document parsing and validation system using Azure OpenAI, tested on 250+ forms and reducing turnaround time by approximately 60% compared to manual processing

University of Arkansas CVIU Lab Fayetteville, AR
Research Assistant *Feb. 2024 – Present*

- Co-authored a physics-informed domain adaptation network (NeurIPS 2025 submission) for aligning synthetic and real 2D material images with color normalization and source-transform modules
- Created a synthetic dataset of 600K microscopy images (≈1M flakes) across eight materials and 40 layer types
- Improved thickness estimation error by 9.1 nm and detection precision by 30%
- Achieved state-of-the-art flake layer classification accuracies up to 93.9%, outperforming prior works by 15%

Projects

ReSearch — *SvelteKit, AWS, Nginx, Python* *Aug. 2024 – May 2025*

- Implemented a high-performance search platform enabling rapid full-text and temporal queries across 50K timestamped webpages spanning months of scraped historical data
- Optimized query processing with vectorized TF-IDF scoring and precomputed indices, reducing latency by 78% and delivering median response times ≤ 400 ms, enabling real-time user interactions on large datasets

Grocify — *Streamlit, Python, LangChain, ViT, OpenAI* *Feb. 2025 – Mar. 2025*

- Built a Streamlit app to automate grocery shopping by converting text recipes or meal images into prefilled Walmart carts, leveraging ViT, Spoonacular API, and OpenAI GPT for precise ingredient extraction
- Delivered a fully automated pipeline with average end-to-end processing time ≤ 8 seconds

San Francisco Crime Classification — *Python, Pandas, Matplotlib, CatBoost* *Sep. 2024 – Dec. 2024*

- Engineered spatio-temporal features from 800K+ SFPD crime reports to classify incidents across 39 categories
- Trained a CatBoost model with stratified 5-fold validation, achieving weighted F1 score of 0.2505
- Ranked in the top 7% (153/2,332) on Kaggle's San Francisco Crime Classification leaderboard