Harshvardhan Nagar

(480) 560-1172 | harshvardhanvn1@gmail.com | LinkedIn | harshvardhannagar.com

Education

Arizona State University **Tempe, Arizona**
Masters of Science in Computer Science,(GPA: 3.83) May 2026
Relevant Coursework: Data Mining, Fundamentals of Statistical Learning and Pattern Recognition, Natural Language Processing, Data Processing at Scale, Knowledge Representation & Reasoning, Software Security.

Pandit Deendayal Energy University **Gandhinagar, India**
B.Tech in Information and Communication Technology,(CPI: 9.47/10) May 2024
Relevant Coursework: Data Structures and Algorithm, Operating System, Object Oriented Concepts & Programming, Database Management Systems, Machine Learning, AI Systems, Deep Reinforcement Learning, Cloud Computing, Information Security, Embedded Systems, Computer Communication & Networking.

Technical Skills

Programming and Scripting: Python, C/C++, Java, SQL, Bash, HTML, CSS, JavaScript, TypeScript, React.js, Node.js
AI/ML Libraries: Pytorch, TensorFlow, Pandas, NumPy, OpenCV, Scikit-Learn, Matplotlib, Seaborn, MMDetection.
Cloud & DevOps: Databricks, AWS (EC2, S3), GCP; Docker, Kubernetes, CI/CD, Github.
Databases & Warehousing PostgreSQL, MongoDB, Snowflake, Redshift, Athena, BigQuery. Apache Kafka, Apache Airflow, Apache Spark (PySpark); Hadoop (Hive, HDFS); Power BI, MS Excel.
Model Architectures: CNNs; Vision Transformers; Language Transformers; Sequence Models; Detection/Segmentation; Generative (VAEs, GANs, Diffusion); Multimodal.

Experience

Data Analytics, Simulation & IoT, AMI Intern **May 2025 - Present**
MAGNA International Troy, Michigan
- Cut camera setup time from 3–4 weeks to less than 1 day by configuring GigE industrial cameras with ROS2 and SDKs, and creating instructional videos and confluence documentation that streamlined deployment across plants.
- Engineered and deployed a custom ROS2 node using camera's HTTP API SDK and Python wrapper, achieving 80% cost savings - earning accolades from senior leadership; followed software development practices and pushed code to GitHub for integration with the AI Vision Platform.
- Aided plant engineers in installing and configuring 24 industrial cameras to initiate data acquisition; selected as 1 of 3 interns supporting deployment of image collection application on plant's virtual machines to enable AI vision platform.

AI/ML Intern **January 2024 - April 2024**
Brandbook Studio Orissa, India
- Expanded an image corpus +19% via targeted web scraping; used Selenium to crawl and BeautifulSoup to parse/collect assets, then applied Copilot + LangChain to generate dense, structured captions for training-ready pairs.
- Built an automated parsing/validation pipeline; applied YOLO for element detection and ResNet18/50 for attribute tags; gated pairs with CLIPScore + BLIP-VQA to verify prompt–image alignment and filter noisy samples prior to training.
- Designed text-to-image generation stack; configured PyTorch Diffusers with ControlNet + LoRA for layout/attribute control, templated prompts, and versioned configs - reproducible fine-tuning path that scales from 2000 images.

Data Analyst Intern **July 2023 - December 2023**
Roopa Screen Private Limited Ahmedabad, India
- Automated preprocessing in Python (Pandas), SQL, and Excel; cut manual review 14 hrs/week by standardizing cleaning steps and running cross-dept EDA to build cleaner, more reliable data pipelines.
- Derived product insights by testing Logistic/Linear Regression, Decision Trees, and Clustering on real data; validated with 3 cross-functional teams and drove feature recommendations.
- Built interactive Power BI dashboards (transactions, AOV, CSAT); surfaced KPIs and alerts that contributed to a 27% revenue lift through faster, data-driven decisions.

Projects

Secure-RAG
- Architected a secure, auditable RAG; hit micro-F1 0.971 (P 0.995/R 0.947) on synthetic PII; used Presidio+spaCy redaction, FastAPI APIs, Postgres 16+pgvector ANN search; shipped via Docker Compose with reproducibility.
- Elevated retrieval to Recall@5=0.67 on BEIR; used sentence-transformers (all-mpnet-base-v2, 768-D) embeddings and tuned pgvector HNSW/IVFFlat in Postgres 16; logged Recall@K to DB for ongoing drift monitoring.
- Enforced per-document ACLs and email-scoped auth; instrumented retrieval tracing (each query+result logged); shipped 5 security/eval features end-to-end; delivered React+Vite+Tailwind UI and Docker dev-up.

Leverage LLMs for Potential SQL Joins
- Developed an evaluation pipeline for join prediction robustness in text-to-SQL, validating LLM outputs against Spider dataset ground truth and mining 249 structured failure cases (76.7% false positives, 23.3% false negatives).
- Built schema-aware SQL parsing and JSON-based prompting workflow, combining sqlglot AST parsing, Gemini API, and schema metadata extraction to ensure reliable and explainable join comparisons.
- Automated error analysis and visualization, implementing restart-safe CSV logging and Matplotlib reports to quantify error distributions across single- vs. multi-table joins.