Dhruv Desai

732-397-5463 | desai.dhruv.1611@gmail.com | linkedin.com/in/dhruv1116/ | github.com/1611Dhruv

EDUCATION

University of Wisconsin - Madison *at Madison, WI* — *GPA - 4.0/4.0*
Bachelor of Science in Computer Science (Honors) & Math for Programming and Computing — Sep 2022 - May 2026
Business, Certificate — Sep 2024 - May 2026

Relevant Coursework — *Honors in Algorithms, Distributed Systems, Operating Systems*
Database Management Systems, Networking, Programming Languages and Compilers
Machine Organization, Artificial Intelligence, Big Data Systems
Game Theory, User Interfaces, Discrete Mathematics, Linear Algebra
Probability Theory, Stochastic Processes, Linear Optimization

EXPERIENCE

Software Engineering Intern *at Commvault, Hybrid* — May 2025 – Present
- Developed gRPC framework in C# for node communication, increasing multipart-upload throughput by ∼30%
- Ensured cache consistency across distributed nodes for multipart uploads, improving reliability and fault tolerance
- Managed build tool reliability, resolving 8+ critical incidents via server log analysis and root cause remediation
- Integrated LLM-based root cause analysis to automate anomaly detection and streamline distributed debugging
- Enhanced React frontend UX through QoL improvements & efficient navigation of a large enterprise codebase
- Built cipher detection model with 80% accuracy & 0.75 F1-score using XGBoost, CNNs, and LightGBM
- Engineered 107 NIST-based probabilistic features in Apache Spark to strengthen cipher classification accuracy
- Integrated C++ extractors with ONNX inference pipeline, achieving 270ms latency for 40KB ciphertexts

Research Assistant *at University of Wisconsin–Madison, Madison, WI* — Jan 2025 – Present
- Developed a LLM-based code evolution framework to enhance Linux PageCache policies
- Extended framework to Linux page compaction, leveraging eBPF and kfuncs for dynamic policy injection
- Comparing Evolve with traditional ML-based adaptive caching frameworks using handcrafted feature models
- Benchmarked evolved policies on Filebench, SPECsfs, and FIO, achieving up to 25% higher throughput

Course Assistant – OS, AI *at University of Wisconsin–Madison, Madison, WI* — Sep 2023 – May 2025
- Mentored 600+ students in OS and AI, resolving 98% of conceptual and implementation-level coding issues
- Taught core OS and AI topics including concurrency, virtualization, CNNs, clustering, and RL
- Authored a GDB-based concurrent debugging guide for kernel-space issues, adopted across 2 semesters by TAs
- Led 10–15 student system design sessions on Copy-on-Write, huge pages, RAID, and lock-free buffer architectures
- Automated grading workflows via GitLab scripts and project setup, saving 10+ TA hours weekly

Software Engineering Intern *at Next Play Games, Lake Orion, MI* — May 2024 – Aug 2024
- Achieved an 80% reduction in API calls by implementing a memoized in-memory caching mechanism
- Benchmarked caching strategies quantifying per-user load reduction from 400+ API calls to 30 calls per page
- Implemented a decoupled architecture to ensure backend changes did not impact frontend components
- Developed a community-driven fantasy baseball app with React, TypeScript, and PostgreSQL
- Led a startup team of 4 interns, ensuring seamless project execution via effective work division and team meetings

PROJECTS

Office Hours Queue | *LLMs, Redis, ChromaDB, Docker, Microsoft AI Foundry* — Present
- Engineered real-time ticketing infrastructure using Redis caching to maintain state and minimize latency
- Architected an ML pipeline using ChromaDB to parse multimodal data, enhancing instructor observability
- Built an LLM-assisted office hours queueing system that cut TA wait times by 80% (40 min → 10 min)
- Led infrastructure design and deployment for a 4-person team, coordinating progress and information flow

MarketPulse – Real-Time Marketplace Intelligence | *Kafka, Flink, Python, FastAPI, Airflow* — Present
- Built a real-time ML system simulating an eBay-style marketplace for a high-demand GPU (RTX 59000)
- Processed auction and bid events with Kafka and Flink to generate low-latency ML features
- Served price prediction and fraud models via production-style inference APIs
- Automated retraining and monitoring with Airflow, tracking drift and system health

TECHNICAL SKILLS

Programming Languages: C, C++, C#, Go, Python, Java, JavaScript, TypeScript
Cloud Platforms: Azure, AWS, GCP, Kubernetes, Docker, Serverless, PubSub, Redis, S3, Azure Blob
Cloud AI & APIs: Microsoft AI Foundry, OpenAI API, Azure Cognitive Services
Web & API Development: Node.js, Express, FastAPI, Flask, Django, React, GraphQL, Auth0, HTML/CSS
Databases & Infrastructure: PostgreSQL, MySQL, MongoDB, ChromaDB, Prisma, Terraform, CI/CD
Developer Tools: Visual Studio, Git, FUSE, GDB, Valgrind, PyCharm, eBPF, Bash, Maven, Make