ADITYA AGARWAL (He/Him/His)
Jersey City, NJ | adiagarwal1509@gmail.com | www.adityaagarwal.me
LinkedIn: @adityaagarwal1999 | GitHub: @adiagarwalrock

TECHNICAL SKILLS & CORE COMPETENCIES

Languages & Frameworks	Python, Django, Flask, Fast-API, Java, JavaScript, PyTorch, CUDA, Celery
Generative AI & MLOps	Generative AI, LLMs, LLM agents, MLOps, Data Science, Evaluation Frameworks, RAG Pipelines, Agentic Systems, BERT, NLP, Vector Databases, LangChain, Llama-Index
Security & Compliance	Data Governance, Automated Monitoring, Security Controls, Content Analysis, Threat Detection
Cloud & Infrastructure	AWS, GCP, Docker, Terraform, System Design, High Availability
Data & Distributed Systems	MongoDB, MySQL, PostgreSQL, ETL Pipelines, Real-time Processing, Event-Driven Architecture, Microservices
Development Practices	Continuous Deployment, Software Engineering, SDLC (Development Life Cycle), Computer Architecture, Decision-making, Git, DevOps, REST APIs, Agile, IaC

PROFESSIONAL EXPERIENCE

Founding Software Engineer | Soopra.ai, San Francisco   **August 2024 - Present**
- Led the **system design** and implementation of **Generative AI** workflows using **LLM agents**, resulting in a 35% boost in user retention.
- Established **MLOps** best practices for deploying large-scale **LLMs**, including automated **evaluation frameworks** to monitor model drift and hallucination rates.
- Built resilient data processing systems creating Agent-first social media platform with high availability and distributed computing principles.
- Applied advanced **Data Science** methodologies to optimize recommendation engines, aligning technical output with executive **decision-making** goals.
- Optimized **inference latency** and token consumption for conversational AI, achieving a 20% reduction in user onboarding time through automated validation scripts.

Software Developer Intern | Soopra.ai, San Francisco   **June 2023 - May 2024**
- Engineered Hybrid-Search with Keyword and Semantic matching on multi-source ingestion pipeline with recency-biased re-ranking, achieving 95% embedding accuracy.
- Designed event-driven chat systems utilizing domain-driven architecture to dynamically determine conversation flow, increasing user engagement by 25% through intelligent response routing
- Optimized background task processing using **Celery** and Redis for real-time social media ingestion, cutting data loading latency by 15%.

Full-Stack Developer | SA Consultant, Bangalore   **October 2021 - June 2022**
- Managed the full **Software Development Life Cycle (SDLC)** for enterprise web applications, applying **System Design** patterns and **Software Engineering** best practices.
- Implemented billing and monitoring systems with PostgreSQL backend, enhancing project documentation accuracy and providing audit trails for compliance
- Deployed cloud-native solution on Heroku with RDS integration, ensuring high availability, data security, and seamless scalability for distributed teams

EDUCATION

Master of Science in Machine Learning   **September 2022 - May 2024**
Stevens Institute of Technology, NJ

Bachelor of Engineering in Computer Science   **August 2017 - July 2021**
Alliance University, Bangalore

KEY PROJECTS

SOAR – Managed RAG System | Product Page
- Built an AI framework using Django, Llama-Index, and Qdrant to turn unstructured documents into a semantic, searchable knowledge graph with low-latency vector search.
- Developed customization tools for live prompt engineering, model swapping (Claude, GPT-4, etc.), and parameter tuning to optimize AI outputs.
- Productized solutions such as embedded widgets & REST APIs with automated PDF/CSV reports, Slack alerts, and ticket creation pipelines.
- Deployed on GCP with Docker/Kubernetes, enabling auto-scaling, security controls, and a roadmap for multimodal (image/table) support.

Toxic Span Detection & Content Security   **March 2023 - April 2023**
- Developed AI-powered security solution using BERT and CRF models for automated toxic content detection in online discussions, enabling proactive threat identification and semi-automated moderation
- Implemented ML-based governance controls highlighting malicious spans within content for enhanced attribute and security monitoring

critiQs - Secure Review Platform | Live Demo   **February 2023 - May 2023**
- Built cloud-native web application with integrated TMDB database containing 1000+ entries, implementing secure user authentication and recommendation algorithms
- Designed scalable architecture with REST APIs and database optimization, enabling high-performance content delivery and user engagement analytics

Insight - Enterprise Project Management Platform   **October 2021 - June 2022**
- Led development of Django-based platform with comprehensive project monitoring, billing estimates, and automated reporting capabilities for enhanced operational efficiency
- Deployed production system on Heroku with AWS RDS, implementing infrastructure-as-code principles for reliable data management and security compliance

Fantastic Computing Machine - ML Platform   **December 2020 - June 2021**
- Created SaaS platform on Flask framework for dynamic ML model deployment, enabling secure model sharing and collaborative development with enhanced deployment efficiency

CERTIFICATIONS & PUBLICATIONS
- Neural Networks and Deep Learning - Coursera (deeplearning.ai) | April 2020
- Published Research: "*Text Mining Approach Based on TF-IDF and SVM for Text Classification*" in *Latest Innovation for Future Education 2021* | ISBN: 978-81-947019-0-3 | January 2021