Aditya Kulkarni

Dallas, TX | adityavkulkarni0999@gmail.com | +1 945-527-5186 | linkedin.com/in/adityavkulkarni | adityavkulkarni.github.io

EDUCATION

University of Texas at Dallas   Dallas, TX
Master of Science, Computer Science   *August 2023 – December 2025*
Independent Research: Visual tracking of aquatic animals using ML-controlled underwater robots

Savitribai Phule Pune University   Pune, India
Bachelor of Engineering, Computer Science   *August 2017 – May 2021*

SKILLS

Programming: Python, JavaScript, SQL, TypeScript
ML/AI: PyTorch, TensorFlow, Transformers, LangChain, LlamaIndex, Scikit-learn, OpenCV
Frameworks: React, Angular, Node.js, Flask, Django, FastAPI
Cloud & Infrastructure: AWS (EC2, S3, Lambda, RDS, CloudWatch, EKS), Docker, Kubernetes
Data & Streaming: PostgreSQL, MongoDB, Neo4j, Apache Kafka, Apache Spark, Redis, ElasticSearch
DevOps & Monitoring: Jenkins, MLflow, Grafana, Prometheus, ELK Stack

WORK EXPERIENCE

University of Texas System, University Lands   Dallas, TX(Remote)
Machine Learning Intern   *June 2025 – Present*
- Architected and deployed a production-grade **Agentic RAG** system processing **200+ reports monthly** with **85%** accuracy in metadata extraction.
- Built **FastAPI** microservice with **Neo4j** integration, achieving **sub-3s** response times for complex legal document queries across **10GB+** knowledge base.
- Integrated seamlessly with existing Angular-based portal, enabling real-time contract analysis and reducing manual review time by **60%**.
- Implemented automated CI/CD pipeline using **Azure DevOps** and **Docker** with zero-downtime deployments.
- Designed fault-tolerant architecture with **PostgreSQL clustering**, achieving 99.9% uptime.

Boehringer Ingelheim   Ridgefield, CT
Data Science Intern   *January 2025 – May 2025*
- Developed and optimized LLM frameworks (GPT-4o, Claude Sonnet 3.5) for vendor classification through webpage analysis.
- Increasing classification accuracy by **8%** by applying advanced **Chain-of-Thought prompting**.
- Built **Streamlit** applications with **multi-processing**, reducing processing time from 3 minutes to 20 seconds per request.
- Designed evaluation pipelines using **Argilla** and **DeepEval**, creating custom dashboards for LLM performance monitoring.
- Led **MLOps** implementation for production deployment on **OpenShift**, integrating **CI/CD** pipelines and enhancing application reliability.

PubMatic Inc.   Pune, India
Software Development Engineer in Test-1   *April 2021 – June 2023*
- Spearheaded end-to-end testing for multiple high-priority, high-revenue projects, ensuring robust performance and reliability under heavy client usage.

PROJECTS

Agentic Parser for PDFs - github.com/adityavkulkarni/agentic-pdf-rag   June 2025 – Present
- Built a Python library enabling advanced **PDF parsing**, **intelligent chunking**, and **visual element extraction** for documents.
- Implemented **dual retrieval strategies** (vector + graph search) achieving **85%** precision in context-aware querying across document corpus.
- Optimized processing pipeline reducing document parsing time by **65%** through async processing and intelligent caching mechanisms .

Realtime Reddit News Analysis - github.com/adityavkulkarni/media-analyzer   May 2024 – August 2024
- Migrated architecture to **AWS serverless** using **Lambda** and **SQS**.
- Built streaming data pipeline with Kafka and Spark, achieving real-time classification with **< 2-second latency** for news sentiment analysis.
- Developed news classification and sentiment analysis models using fine-tuned **DistilBERT** and **SparkNLP ClassifierDL** with classification accuracy of **83%** and sentiment analysis accuracy of **91%**.

InvestAid: An AI-powered Investment Dashboard - github.com/sauravdosi/investaid   October 2023 – March 2024
- Leveraged Spacy models to analyze sentiment and topics in over 50,000 social media posts and financial news articles, achieving a classification accuracy of 85% and identifying key topics relevant to stock performance.
- Integrated historical data analysis to generate data-driven stock position recommendations for top 3 performing stocks.

CERTIFICATIONS

IBM Data Science Certification