

Aditya Kulkarni

Dallas, TX | adityavkulkarni0999@gmail.com | +1 945-527-5186 | [linkedin.com/in/adityavkulkarni](https://www.linkedin.com/in/adityavkulkarni) | adityavkulkarni.github.io

EDUCATION

University of Texas at Dallas

Master of Science, Computer Science

Independent Research: Visual tracking of aquatic animals using ML-controlled underwater robots

Savitribai Phule Pune University

Bachelor of Engineering, Computer Science

Dallas, TX

August 2023 – December 2025

Pune, India

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

Machine Learning Intern

June 2025 – Present

University of Texas System, University Lands

Dallas, TX(Remote)

- Architected and deployed a scalable Agentic **RAG** system processing 200+ legal reports monthly with 85% accuracy in metadata extraction, reducing manual review effort by 60% and enabling compliance analysis.
- Designed a modular system supporting Azure OpenAI, Ollama, and Qwen3 Models.
- Built **FastAPI** microservice with **Neo4j** integration, achieving sub-800ms response times for complex legal document queries across 10GB+ knowledge base.
- Implemented multilevel retrieval combining semantic chunking, query expansion, and knowledge graph traversal.
- Integrated seamlessly with existing **Angular**-based portal, enabling real-time contract analysis and reducing manual review time by 60%.
- Delivered chatbot, portal integration, and daily contract compliance alerts reducing manual work efforts for legal analysts.

Data Science Intern

January 2025 – May 2025

Boehringer Ingelheim

Ridgefield, CT

- Developed and optimized LLM frameworks for vendor classification through webpage analysis reducing outsourcing.
- 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 **DeepEval**, creating custom dashboards to monitor LLM performance and model drift in real time.
- Led MLOps implementation for production deployment on **OpenShift**, integrating CI/CD pipelines and enhancing application reliability.

Graduate Research Assistant

May 2024 – December 2024

University of Texas at Dallas

Dallas, TX

- Engineered computer vision system for underwater robotics using YOLO and ResNet50, achieving 92% detection accuracy on 10,000+ marine animal images.
- Finetuned CodeLlama model using LoRA and quantization techniques, improving software engineering task accuracy to 89%.

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 for **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**, **S3** 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](#)