

# Aditya Kulkarni

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

## SUMMARY

Results-driven Machine Learning Engineer with experience creating and implementing AI solutions in legal, healthcare, and finance. Proven track record of improving business processes and outcomes by developing practical, reliable systems. Skilled at leading projects from concept to completion and collaborating with diverse teams to deliver measurable results in real-world settings

## EDUCATION

### University of Texas at Dallas

Master of Science, Computer Science. **GPA: 3.74/4**

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

Dallas, TX  
August 2023 – December 2025

### Savitribai Phule Pune University

Bachelor of Engineering, Computer Science. **GPA: 9.12/10**

Pune, India  
August 2017 – May 2021

## SKILLS

**Programming:** Python (proficient), C++ (proficient), SQL

**ML/AI:** Deep Learning, NLP, Computer Vision, LLMs (Transformers, PyTorch, HuggingFace)

**Frameworks:** LangChain, LlamaIndex, Spark (SparkSQL), DeepEval, Argilla

**Tools:** Docker, Kubernetes, Jenkins, MLFlow, Flask, Databricks, ELK Stack, Linux, Git, AWS (S3, EC2, Lambda), Kafka

**Databases:** MongoDB, MySQL, Neo4j, PostgreSQL, HDFS

## WORK EXPERIENCE

### University of Texas System, University Lands

*Machine Learning Intern*

Dallas, TX(Remote)  
June 2025 – Present

*Tools: GPT4o, Langchain, Neo4j, OpenAI API, Qwen3, PostgreSQL, Streamlit, Pytorch, HuggingFace Transformers*

- Built an advanced **Agentic Graph-based RAG** pipeline to parse, structure, and store contracts, leases, and sequential amendments as interconnected knowledge graphs, enabling relationship analysis and efficient retrieval.
- Automated **query classification** and **tool selection** using **agentic workflows**, enabling intelligent selection between **vector search, graph traversal, and hybrid approaches** based on the user's query type and context.
- Enabled **dynamic, recursive retrieval** and robust error handling through **agent-driven feedback loops**, ensuring accurate extraction of contract metadata, parties, dates, monetary values, and amendment relationships for legal and compliance use case.

### Boehringer Ingelheim

*Data Science Intern*

Ridgefield, CT  
January 2025 – May 2025

- Developed and optimized LLM frameworks (GPT-4o, Claude Sonnet 3.5) for vendor classification and webpage analysis, increasing classification **accuracy** by **8%** by applying advanced prompt engineering (**Chain-of-Thought**) and parameter-efficient fine-tuning (**PEFT**).
- Built **Streamlit** applications with multi-processing, reducing processing time by 30%.
- 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.

### University of Texas at Dallas

*Graduate Research Assistant*

Dallas, TX  
May 2024 – December 2024

- Engineered a specialized **object detection system** for aquatic animals using **ResNet50** and **YOLO** architectures.
- Analyzed over 25 articles on state-of-the-art underwater robotics and vision tracking systems and presented a detailed review.
- Partnered with senior researchers at HBS Lab to advance **ML vision** capabilities for **underwater robot navigation**.
- Spearheaded research on **fine-tuning CodeLlama** for software engineering tasks, using **PEFT** methods like **LORA** and **quantization** achieving a **84%** accuracy.

### PubMatic Inc. Pune, India

*Software Development Engineer in Test-1*

April 2021 – June 2023

## PROJECTS

### Agentic Parser for PDFs - [github.com/adityavkulkarni/agentic-pdf-rag](https://github.com/adityavkulkarni/agentic-pdf-rag)

June 2025 – Present

- Built a Python library enabling AI-powered PDF parsing, intelligent chunking, and visual element extraction for documents.
- Implemented dual retrieval strategies and vector search, supporting precise, context-aware querying and knowledge base creation.
- Delivered flexible integration as a library or API, supporting legal, research, and technical document workflows at scale.

### Realtime Reddit News Analysis - [github.com/adityavkulkarni/media-analyzer](https://github.com/adityavkulkarni/media-analyzer)

May 2024 – August 2024

- Engineered a real-time news classification and interpretation dashboard using **Apache Kafka, Spark Structured Streaming** and **Kibana** dashboard for visualizing real-time results using **ELK stack** with capability of processing **50** articles per minute.
- Developed news **classification** and **sentiment analysis** models using fine-tuned **DistilBERT** and **SparkNLP ClassifierDL** with bias classification accuracy of 83% and sentiment analysis accuracy of 91%.

### InvestAid: An AI-powered Investment Dashboard - [github.com/sauravdosi/investaid](https://github.com/sauravdosi/investaid)

October 2023 – March 2024

- Designed an AI-powered **investment advisory** system leveraging **NLP** and **aspect-based sentiment detection** techniques.
- 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 10 key topics relevant to financial markets.
- Integrated historical **data analysis** to generate data-driven stock position **recommendations** for top 3 performing stocks.

## CERTIFICATIONS

[IBM Data Science Certification](#)