ABHISHEK DABAS

adabhishekdabas@gmail.com | linkedin.com/in/adabhishek | github.com/abhishekdabas31
AI Scientist | Agents, LLMs, RAG, Ethical AI | Driving Efficiency & Innovation

EDUCATION

Master of Science in Information Systems **Northeastern University**, Boston, MA
Relevant Courses: Artificial Intelligence, Data Science Engineering, Reinforcement Learning, Machine learning & NLP
Experience: President of the largest graduate student organization (AI Skunkworks), RA & TA at COE

TECHNICAL SKILLS

ML Engineering & MLOps: Expertise in end-to-end ML orchestration (research to production), AWS SageMaker (training & deployment), Bedrock, Azure Open AI, Lambda, Step Functions, monitoring frameworks, automated testing pipelines
Generative AI: Advanced Prompt engineering, LLM fine-tuning/deployment, and agentic frameworks (LangChain)
Machine Learning: TensorFlow, PyTorch, LIME/SHAP, Regression, time-series modeling, A/B testing in python
Domain Expertise: Financial data processing, unstructured text analysis, document processing

WORK EXPERIENCE

Data Science, Manager - Fidelity Investments AI COE **Feb 2022 – Present**
***LLM Agent for Back-Office Automation*:** Architected and deployed an enterprise LLM-powered agent framework using Claude 3.0 and Llama 3, focusing on document processing and knowledge retrieval automation. Implemented sophisticated prompt engineering with continuous feedback loops from back-office associates for iterative improvements. Designed an evaluation framework to measure agent effectiveness and establish performance benchmarks.
Ethical AI Research: Led the development and implementation of ethical AI practices within my BU, including HELM-based metrics and robust LLM benchmarking frameworks, to ensure model fairness, transparency, and alignment with product requirements. Presented findings to senior leadership, driving the adoption of responsible AI practices.
LLM Evaluation & Benchmarking: Built large, high-quality datasets for CV, classification & QA tasks using Label Studio, employing rigorous annotation guidelines and quality checks to ensure data accuracy & reliability for robust LLM evaluation.
RAG for Document Search & Extraction:Led end-to-end development of a RAG using the lamaindex framework and Claude 3, from A/B testing framework design to production deployment, achieving 30% efficiency gain (equivalent to 2-4 FTE savings) processing 100k+ documents monthly. Architected scalable AWS infrastructure using Step Functions and CloudWatch for real-time performance tracking.
Intelligent Document Processing: Led development of document automation platform using YOLO and LayoutLM, achieving 50% processing time reduction through rigorous A/B testing of processing workflows. Designed evaluation framework with performance monitoring, data quality gates, & dynamic retraining triggers ensuring production reliability.
Mentorship & Collaboration: Actively mentored Co-Ops and collaborated with cross-functional teams to align technical solutions with business strategy, contributing to more effective stakeholder management.

Machine Learning Research Assistant, Northeastern University **July 2021 – Dec 2021**
***MLOps & AutoML Development*:** Led development of custom TensorFlow Extended (TFX) components for automated feature engineering and model validation, integrating into existing MLOps framework. Implemented distributed training pipelines and automated hyperparameter tuning, ensuring scalability and reproducibility for production ML workflows.

Data Science, Co-Op - Fidelity Investments AI COE **Jan 2021 – July 2021**
Financial NLP & Market Trend Analysis: Developed an end-to-end pipeline for an NLP-based fund market trends analysis, achieving 91% accuracy with fine-tuned BERT. Defined long-term roadmap for automated market intelligence, building partnership with trading desk to continuously optimize model parameters based on real-world performance. Deployed this as a scalable solution on AWS (EC2, S3) with custom-built UI for portfolio managers.
Ethical AI Research: Developed bias detection framework for NLP models using counterfactual data augmentation and statistical significance testing. Built automated fairness evaluation pipeline with custom metrics for gender bias across multiple languages, achieving 40% reduction in model bias while maintaining performance.