Abhinav Aditya

Boston, MA | aditya.ab@northeastern.edu | +1 (857) 351-5712 | GitHub | LinkedIn

EDUCATION

Northeastern University, Boston, MA **DECEMBER 2026**

Master of Science, Data Analytics Engineering **GPA 3.92/4**

Relevant Courses: Deep Learning for AI, Applied Generative AI, MLOps, Data Management for Analytics, Data Mining in Engineering

Manipal University Jaipur, Jaipur, India **JUNE 2022**

Bachelor of Technology, Computer and Communication Engineering **CGPA 8.13/10**

Relevant Courses: Design & Analysis of Algorithms and DAA Lab, Advanced Data Structure, Machine Learning, ML Lab

PROFESSIONAL EXPERIENCE

Agentdesks Pvt. Ltd., Bengaluru, India (subsidiary of **Radius Agent, San Francisco**) **JUNE 2022 – JUNE 2024**

Software Development Engineer

- Led end-to-end development of a generative AI chatbot by fine-tuning OpenAI's **gpt-3.5-turbo** model on **500+ documents**, reducing average response time from **5+ seconds** to less than **3 seconds** using caching
- Optimised **RAG retrieval** by benchmarking embedding dimensions, upgrading from **384 to 1024 after K=3 evaluation** to improve semantic accuracy
- Identified critical user **drop-off points** using **Heap funnels**, then developed a new homepage (NextJS) based on the findings, resulting in a **19% increase** in sign-ups in 3 months
- Architected a CI/CD pipeline using GitHub Actions with **3 parallel jobs** and **esbuild**; implemented automated unit tests and reported **78.3% code coverage** via Codecov, contributing to a **57% reduction** in build time
- Led full-stack revamp of CRM tool serving **1,000+ agents**, **migrating** from VanillaJS to **ReactJS** with **Redux**, integrating 10+ **RESTful APIs**, and improving **lead conversion rate by 17%**
- Collaborated with a **12-person cross-platform team** to develop 'Radius Rooms,' a WebRTC solution using the Agora SDK; supported **100 concurrent users** per room by implementing host, speaker, listener, and screen sharing features, resulting in **500+ monthly active users** and a **6% MoM increase** in user engagement

ACADEMIC PROJECTS

CiteConnect (presented in GoogleExpo at Cambridge, Boston) **SEPTEMBER 2025 - PRESENT**

Developing a scalable, end-to-end recommendation system to help researchers discover relevant academic literature through semantic search and citation graph analysis

- Engineered production MLOps pipeline with automated CI/CD using GitHub Actions, deploying dual SentenceTransformer models (MiniLM-L6-v2, SPECTER2) to Google Cloud Run, with real-time model performance **monitoring and automated retraining triggers**
- Designed an adaptive user personalization system with multi-stage user binning and a 6-factor scoring algorithm (semantic similarity, citations, recency, ground truth, reading level, diversity) that dynamically learns user preferences through **click behavior analysis** and **automated weight optimization**
- Open sourcing the platform; developing **funding prediction** features to **assess grant success probability** through citation analysis

BOTBOT - Business Optimisation Transformer BOT **FEBRUARY 2025 – MARCH 2025**

Developed a multi-agent chatbot to automate business intelligence insights and visualisations from user-uploaded CSVs using LangChain

- Engineered a multi-agent system with LangChain, using a central **Activation Agent** to **dynamically route** user queries to specialised Data Cleaning, Analysis, and Plotting agents
- Implemented a Pinecone vector store to provide conversational memory, **caching previous queries** to **reduce redundant LLM calls** and **optimise response time** with a similarity threshold of 0.8
- **Benchmarked LLM performance** (GPT-4o vs. GPT-4o-mini vs. Gemini) for analytics tasks and **automated insight generation** on any raw sales dataset, identifying **GPT-4o** as the top performer for **graph generation** and insight extraction

Prediction of Insurance Claim **JANUARY 2025 – APRIL 2025**

Built predictive models to assess flood claim approval likelihood by merging property data with storm indicators

- Integrated **FEMA's NFIP claims data with NOAA storm records** to create a comprehensive merged dataset, cross-referencing building features with storm-related events
- Trained and tuned Logistic Regression, Random Forest, Gradient Boosting, and MLP models using GridSearchCV and SMOTE, achieving **0.96 ROC AUC**
- Visualised model comparison using bar plots and SHAP; interpreted **storm-driven vs structural feature** importance to guide disaster risk pricing

SKILLS

Programming Languages: Python, SQL, C++, JavaScript, HTML, CSS

Tools: Git, Postman, Tableau, Jupyter Notebook, VSC, Docker, Heap, AWS, GCP, n8n, ELK, Airflow, MLflow, Kafka, Spark

Libraries: NumPy, Pandas, Matplotlib, scikit-learn, Langchain, LangGraph, TensorFlow, PyTorch, python_a2a, Seaborn, Plotly, Folium

Database: MongoDB, Pinecone, MySQL, Django, Neo4j