

# Akanksha Ashish Mathpati

akankshamathpati8@gmail.com • Chicago, IL • [LinkedIn](#) • [Website](#) • [GitHub](#)

## EDUCATION

**University of Chicago** Expected, Dec 2025  
Master of Science, Applied Data Science CGPA: 4.0/4.0  
Coursework: Big Data and Cloud Computing, Statistical Models, Machine Learning, Time Series, Natural Language Processing

**Dr. D Y Patil's Ramrao Adik Institute of Technology** Oct 2020 - Jun 2024  
Bachelor of Technology in Computer Engineering CGPA: 9.99/10  
Coursework: Database Management Systems, Data Warehousing & Mining, Artificial Intelligence, Big Data Analytics

## EXPERIENCE

**Blue AI Labs (Nevara)** Chicago, IL  
**Machine Learning Engineer** Jun 2025–present

- Co-led a **Python/Flask + Supabase** application that uses GPT-based LLMs, prompt engineering, and RAG to turn sales-call transcripts into **personalized coaching feedback**—achieving **82% alignment** with coach-written reports.
- Built **end-to-end machine learning pipelines** (real-time feedback loops and performance dashboards) to monitor **LLM output quality** and refine **prompts**, enhancing the contextual relevance and usefulness of AI-generated insights.

**FA ITS - Chief Technology Office, University of Chicago** Chicago, IL  
**Machine Learning Engineer** Nov 2024–present

- Implemented features in **Phoenix AI**, a **ChatGPT-like conversational AI system** tailored for university services using **Python, Git, GPT-4o, Docker, Azure OpenAI, Cosmos DB**.
- Develop a **Retrieval-Augmented Generation (RAG) model** for the HR department, streamlining application processing and employee information management using **Python, Langchain, Ollama and Vertex AI**.
- Deployed and optimized **ML-powered applications**, including LibreChat, by integrating machine learning solutions into existing business intelligence frameworks enhancing accessibility and decision-making.

**Bhabha Atomic Research Centre** Mumbai, India  
**Project Intern** May 2023 – July 2023

- Engineered a robust **Python + Flask based solution** to automate PDF data extraction and form-filling tasks; deployed a scalable web service reducing manual processing time by **80%** and improved **operational efficiency by 65%**.

## PROJECTS

**Law GPT** Aug 2023 – Dec 2023

- Created an **Large Language Model (LLM) Powered Chatbot** used to simplify Indian legal information, translating legal jargon into plain language, and **integrating voice assistance** for improved accessibility.
- Leveraged **natural language processing (NLP)** to answer queries and improve user experience in legal consultations.

**Stackoverflow NLP Analysis** Oct 2024 – Dec 2024

- Conducted **NLP and sentiment analysis** of StackOverflow posts using **PySpark NLP and Spark ML** to derive insights into user interactions and post trends.
- Utilized big data processing techniques on **Google Cloud Platform** to efficiently handle 30 GB of data, achieving over **95% accuracy on 50 logistic regression models** for NLP analysis of StackOverflow posts.

## SKILLS & PROFICIENCIES

**Programming:** Python (pandas, scikit-learn, seaborn, NumPy, matplotlib, nltk, spaCy, LangChain, PyTorch, TensorFlow), R, SQL, PySpark

**Tools:** GCP, AWS Cloud, Vertex AI, Tableau, Power BI, Docker, Jupyter, Visual Studio Code, GitHub, Microsoft Suite, Notion

**Data Science Techniques:** Data Preprocessing, Feature Engineering, Machine Learning (Linear Regression, Logistic Regression, XGBoost, Random Forest, K-Means Clustering, DBSCAN), Natural Language Processing (Sentiment Analysis), Data Visualization, Large Language Models ((OpenAI, LLaMA, Gemini), RAG

**Soft Skills:** Communication, Leadership, Problem-solving, Attention to Detail, Presentation Skills

## PUBLICATIONS & COPYRIGHTS

**Research Paper on Dynamics of Cyberbullying** Jan 2024

**Author, 14th International Conference on Cloud Computing, Data Science & Engineering**

- Published “**Dynamics of Cyberbullying on Twitter: ML Detection Models and the Catalytic role of Tweets engagement metrics**” in the IEEE conference, indexed in Scopus, exploring **machine learning models** for detecting cyberbullying on Twitter.
- Achieved **88% accuracy** in detecting cyberbullying through tweet **engagement metrics and sentiment analysis**.

**MedsXplore-Unlock Medicinal Knowledge, (Officially copyrighted by Government of India)** Feb 2024

- Built an **AI-powered** medicine identification app using **OCR for text extraction**, providing comprehensive information on drug composition, side effects, and usage with multilingual support.