

Manya Jain

Tempe, AZ | manyajainrkm@gmail.com | +1 480-637-1065 | linkedin.com/in/manya-jainn | github.com/ManyaJainrkm

Education

Arizona State University

Master of Science in Computer Science

GPA: 3.50/4.00

Tempe, AZ

Expected May 2027

Manipal University Jaipur

Bachelor of Technology in Computer Science Engineering

India

Aug 2021 – May 2025

Research: Co-Author, *Mathematical Models of AI and Its Legal Boundaries* - formalized AI bias/ethics as quantifiable metrics

Skills

AI/ML & GenAI: LLM evaluation, structured outputs, RAG, NLP, prompt design, agentic workflows, Snowflake, LangChain

Data & ML: pandas, NumPy, scikit-learn, XGBoost, statistical modeling, fairness analysis, Tableau

Languages: Python, TypeScript, JavaScript, SQL, Java, C, C++, Bash

Frameworks & APIs: React, TanStack Start, Node.js, REST APIs, LLM streaming APIs, JSON Schema

Systems & Tools: Linux, Docker, Kubernetes, Jenkins, CI/CD, Git

Databases: PostgreSQL, MySQL, Neo4j, Redis

Experience

Safexpress Private Limited

Software Engineer Intern, GenAI

Gurugram, India

May 2026 – Aug 2026

- Evaluated a Snowflake Cortex-powered GenAI BI Copilot against ground-truth sales data across India's largest B2B logistics network, systematically identifying failure modes in LLM-generated SQL, response formatting, and geographic reasoning
- Found that mandatory data-integrity filters excluding cancelled/void transactions were applied inconsistently, producing confident but incorrect figures; restructured rule enforcement to prevent invalid records from corrupting reported numbers
- Designed and refined orchestration logic, geo-disambiguation rules, and output contracts across skill-file iterations to improve response accuracy and prevent reasoning drift on complex multi-step queries

CIPS-AI Lab, Arizona State University

Student Researcher, Applied AI Systems

Tempe, AZ

Feb 2026 – Present

- Identified and validated bigram and trigram feature extraction for legal NLP classification, enabling the model to recognize multi-word domain-specific signals like 'consent judgment' as semantic units
- Built Python ETL pipelines to clean and structure large volumes of unstructured documents for downstream ML classification, enforcing reproducibility through structured logging across 20+ evaluation runs
- Prototyped Neo4j knowledge graphs linking cases, firms, and statutes to support Cypher-based querying and retrieval-augmented workflows

National Institute of Urban Affairs, Govt. of India

Software Engineering Intern, DevOps & Platform Systems

New Delhi, India

Aug 2024 – May 2025

- Built and maintained Jenkins CI/CD pipelines across 12+ containerized services, improving release reliability by 30% through build validation, configuration checks, and deployment gating
- Reduced incident recovery time by 25% by tracing Linux logs, container states and OAuth failures to isolate root causes and restore service within SLA

Ministry of Housing and Urban Affairs, Govt. of India

Data Analytics Intern

New Delhi, India

Dec 2023 – Feb 2024

- Built Tableau dashboards analyzing PAN-India urban development data across multiple years and government schemes, producing public-facing visualizations and internal reporting deliverables used by ministry stakeholders
- Designed SQL-backed data pipelines and validation workflows to process noisy, heterogeneous datasets, and developed Python-based anomaly detection to flag inconsistencies across multi-source data feeds

Recent Projects

Sage: Verified AI Career-Gap Analyzer

TypeScript, React, TanStack Start, Node.js, OpenAI Responses API, JSON Schema, pdf.js

Jul 2026

- Built and deployed a GPT-5.6 reasoning system that analyzes resumes against arbitrary job descriptions, producing calibrated fit scores, prioritized gaps, and evidence-backed strengths through strict structured outputs
- Engineered deterministic verification for resume citations and recurring-gap claims, with browser-side document parsing, server-side inference, retries, safe error handling, and degraded-result reporting

Setu: AI Mental-Wellness Support Platform

JavaScript, HTML/CSS, Streaming LLM APIs, JSON, GitHub Pages

Jul 2026

- Built and deployed a client-only AI support platform with seven methodology-specific guides, a three-question recommendation flow, and real-time streaming chat using user-provided model API keys
- Designed safety-oriented conversation controls including persona-specific scope boundaries, crisis detection and resource surfacing, crisis-first response protocols, browser-only conversation storage, and user-controlled data deletion

Fairness-Aware Income Classification System

Python, scikit-learn, pandas, XGBoost, Git

Jan 2026 – May 2026

- Built an end-to-end ML pipeline on 32,000+ census records achieving ROC-AUC > 0.85 with Logistic Regression and XGBoost, implementing demographic parity and equalized odds checks to quantify model bias and fairness trade-offs