Anshumaan Garg

India ✉ anshumaan.garg001@gmail.com 📞 +91 9599485421 in anshumaan-garg-1a06bb1ba

Education

The University of Texas at Austin, Texas *Sept 2025 – Present*
Post Graduate Program in Artificial Intelligence and Machine Learning (PGP-AIML)
(Offered in collaboration with Great Learning)

- **Coursework:** Probability, Statistics, Linear Algebra for ML, Machine Learning, Deep Learning, Natural Language Processing, Computer Vision, Reinforcement Learning
- Awarded **Meritorious Scholarship** based on academic and professional profile

Amity University *Sept 2020 – Aug 2024*
BTech in Computer Science and Engineering

- GPA: 8.62/10
- **Coursework:** Theory of Computation, Data Structures, Artificial Intelligence, Databases, Software Project Management, Software Testing
- Awarded **Meritorious Scholarship** for academic excellence

Jaspal Kaur Public School, CBSE Board *Mar 2017 – Mar 2020*
Class XII (PCM with Computer Science) and Class X

- Class XII (Apr 2018 – Mar 2020): 92.6% — Subjects: Mathematics, Physics, Chemistry, C++, English
- Class X (Mar 2017 – Mar 2018): 96% — Subjects: Mathematics, Science, Social Science, English

Experience

AI Software Developer (Part-time + Full-time) *Noida, India*
Opkey *June 2023 – August 2025*

- Built cost-effective and low-latency solutions for ERP automation, streamlining enterprise workflows through AI-driven automation and real-time performance optimization.
- Designed and implemented a **classifier-based architecture** within multi-agent LLM workflows for ERP automation, reducing unnecessary token parsing and achieving a **40x reduction in token-related costs** with a **75% overall decrease in processing expenses**.
- Integrated streaming capabilities into LLM pipelines, achieving an **85% reduction in latency**, which was pivotal for **client demos** and improved solution responsiveness, driving stakeholder interest.
- Built an AI-powered automation pipeline to generate **4,000+ product descriptions**, reducing weeks of manual work to just a few days at a cost of only **$1.25**, supporting the **Accelerator team** and **Director of Product − Workday**.
- Engineered **fake streaming mechanisms** within LLM workflows to simulate responses during development, resulting in an **82% reduction in latency**, accelerating testing cycles and improving developer efficiency.
- Optimized **LLMOps pipeline** using **cloud-based GPU hosting** with **scalable deployment templates**, ensuring cost-effective and low-maintenance AI infrastructure.
- Finetuned BERT model for PII detection and classification with **95% accuracy**.
- **Architected and implemented** a **RAG-enabled backend AI system**, reducing model size from **235B** to **8B parameters**, cutting operational costs by **95.83%** while maintaining accuracy and scalability.
- **Engineered** a **multi-class few-shot orchestrator** with **94% accuracy**, capable of asking clarifying questions and dynamically assigning agents based on use case.
- **Implemented Chain-of-Thought reasoning** for agent-based tool selection using **sequential and dynamic I/O referencing**, enabling accurate and context-aware responses.
- **Developed** an LLM-based summarizer converting raw text into well-structured HTML pages, eliminating the need for parsing JSON or Excel files.
- **Built Generative AI–powered chatbot models** to automatically generate test cases from natural lan-

guage prompts, integrating **test case mining techniques**, **JSON-based tagging**, and **OpenAI fine-tuning** for SaaS-based test case creation, management, and optimization.

Publications and Patents

[1] **Generative AI for Software Test Modelling with a Focus on ERP Software**
Anshumaan Garg, Dolly Sharma
Published: Oct 29, 2023
DOI: 10.1109/ICICCT58359.2023.10466102 ↗

[2] **Yoga Posture Analysis Using Deep Learning**
Anshumaan Garg, Geetika Munjal, Sagar Wadhwa
Published: Mar 15, 2024
DOI: 10.1109/ICICCT58359.2024.10522380 ↗

[3] **AI Insights: Navigating Education News Ethically Through Aggregation and Sentiment Analysis**
Anshumaan Garg, Dolly Sharma
Published: Jul 11, 2025
Wiley Online Library ↗

[4] **Beyond the Stars: AI Applications in Space Technology (Planning and Scheduling)**
Anshumaan Garg, Dolly Sharma
Ready for Publication

[5] **System for Real Time AI based Misconception Detection and Validation**
Anshumaan Garg, Kriti Arora, Atia Naim, Dolly Sharma
Filed: 5/15/2025, Patent Pending
Application No: IND 202511046943

Certifications

[1] **IBM Data Science Professional Certificate**
Issued by IBM — Completed via Coursera
Key Learnings: Data analysis, visualization, Python for data science, and ML fundamentals
Date: Aug 2023
Certificate Link: View Certificate ↗

[2] **Pre-training LLMs**
Issued by Upstage — Completed via DeepLearning.ai
Key Learnings: Transformer Fundamentals, Machine Learning
Date: Aug 2024
Certificate Link: View Certificate ↗

[3] **LLMOps**
Issued by DeepLearning.ai
Key Learnings: Data Processing, Monitoring, Observability
Date: Aug 2024
Certificate Link: View Certificate ↗

[4] **Neo4j and LLM Fundamentals**
Issued by Neo4j
Key Learnings: RAG, Langchain fundamentals
Date: June 2025
Certificate Link: View Certificate ↗

Projects

News Aggregation and Analysis Using web scraping and sentiment analysis   *View Project* ↗

- Developed a web application that displays keyword-based news along with media, descriptions, links, and timestamps. Integrated sentiment analysis using VADER on tweets fetched via the Twitter API, storing results in a database. If the same term is searched again after 7 days, new tweets are retrieved automatically.
- Tools and Languages Used: Python, Problem solving, Django, SQL

Falcon 9 Rocket Launch Analysis   *View Project* ☒

- **Falcon 9 Stage 1 Reusability Prediction:** Developed a predictive model to determine whether Falcon 9's first stage could be reused based on launch location, temperature, and payload mass, achieving **83% prediction accuracy**.
- **Tools and Technologies:** Python, pandas, matplotlib, Seaborn, SQL

Honors & Achievements

[1] **Model United Nations (MUN)**
Recognition: Certificate of Excellence - High Commendation
Issuing Organization: Jaipuria Institute of Management
Date: December 2022

[2] **National Science Olympiad**
Recognition: Bronze Medalist
Issuing Organization: Science Olympiad Foundation— TechFest , IIT Bombay
Date: December 2018

[3] **Inter-School Urban Sustainability Challenge**
Recognition: Certificate of Commendation, 2nd position
Issuing Organization: Maharaja Aggarsain Adarsh Public School
Date: November 2014

Volunteering

[1] **JAX Foundation**
National Volunteer
Contributed to child education initiatives and facilitated training in vocational and soft skills, including public speaking

[2] **Delhi Langar Sewa Society**
Community Volunteer
Assisted in preparing and distributing meals to underprivileged communities, ensuring food accessibility and dignity for all

Technologies

Languages: Python

Frameworks: FastAPI , Flask , Langchain, LLamaIndex

Tools and Technologies: LM-Studio, Postman,Ollama,Redis,Docker,Qdrant,Langfuse