

Jatin Gyass – Generative AI Engineer and Full-Stack Developer Specializing in LLMs and RAG Systems

M.Sc. AI student at IIIT Lucknow, building advanced AI solutions with LangChain, Hugging Face, FastAPI—and collaborating with Saksham Pathak on the groundbreaking FALCON project



Lucknow, Uttar Pradesh Jul 23, 2025 ([IssueWire.com](https://www.issuewire.com)) - In a tech landscape being rapidly reshaped by **Generative AI, Large Language Models (LLMs)**, and intelligent automation, **Jatin Gyass** is emerging as a standout young engineer who bridges the gap between **applied AI research** and **production-grade software systems**. Currently pursuing his **M.Sc. in Artificial Intelligence & Machine Learning (2024–2026)** at the **Indian Institute of Information Technology (IIIT) Lucknow**, Jatin is developing **LLM-powered applications, Retrieval-Augmented Generation (RAG) pipelines, and scalable backend services** that map directly to real-world use cases.

From **Yamuna Nagar, Haryana** to a leading national tech institute, his journey reflects curiosity, persistence, and a commitment to making AI practical, accessible, and trustworthy.

A Curiosity That Became a Career Path

Growing up in Yamuna Nagar, Jatin was the kind of student who took things apart to see how they worked—then asked how software could make them smarter. That early mindset led him to pursue a **B.Sc. in Computer Science (2018–2021) from Kurukshetra University**, where he grounded himself in programming fundamentals, databases, and algorithmic thinking.

Eager to understand the **business and product side of technology**, he pursued advanced management studies (Marketing focus, 2021–2023) at SPS Janta College in Mustafabad—experience that now helps him scope AI solutions with clear user value. Seeking deeper technical mastery, he is now fully immersed in **AI & ML graduate study at IIIT Lucknow**, where transformers, vector search, inference optimization, and AI deployment pipelines are part of daily work.

“For me, Generative AI is not just model output—it’s a collaborative layer between humans and intelligent systems,” says Jatin. “I want to build AI that is useful, secure, and production-ready—not just demo-ready.”

Why Generative AI (GenAI) Matters—and Where It’s Going

Jatin believes GenAI is entering a second wave: **from demos to domain depth**. Early public models showed the creative potential of LLMs; the next wave is all about **context grounding, retrieval, and enterprise deployment**. That’s where **RAG (Retrieval-Augmented Generation)** comes in: instead of trusting a model’s frozen training data, RAG systems retrieve fresh, domain-specific knowledge from vector stores and feed it to an LLM at inference time—improving accuracy, relevance, and trust.

He actively engineers such pipelines using:

- **LangChain / LangGraph** for structured multi-step reasoning flows.
- **Hugging Face Transformers** for embedding, fine-tuning, and lightweight model experimentation.
- **FAISS & ChromaDB** for vector similarity search.
- **Prompt engineering frameworks** to steer models for support, analytics, and verification tasks.

“RAG is the difference between a model that ‘sounds smart’ and a system that ‘knows your data,’” Jatin says. “That’s the future of GenAI in enterprises.”

Technical Capabilities at a Glance

Generative AI & LLMs: LangChain, Hugging Face, LlamaIndex, LangGraph, OpenAI API, Prompt Engineering, RAG (FAISS, ChromaDB), context injection, eval loops.

Machine Learning / MLOps: Scikit-learn, NumPy, Pandas, MLflow for experiment tracking & versioning, dataset preprocessing pipelines.

Backend & APIs: FastAPI (Python), Node.js, Express.js, RESTful service design, JWT auth, modular service layers.

DevOps & Deployment: Docker containers, PM2 process manager, NGINX reverse proxy, AWS EC2/S3 hosting, CI/CD basics, Kubernetes (introductory use).

Frontend Engineering: React.js, JavaScript, Tailwind CSS, Bootstrap, HTML/CSS rapid prototyping.

Data & Storage: MySQL, MongoDB, embedding stores (ChromaDB), FAISS vector indexes.

Programming Languages: Python, JavaScript, SQL, Java.

Tools: VS Code, Jupyter, Git/GitHub, Postman, Hugging Face Hub, Streamlit for interactive demos.

Extended Academic Profile

M.Sc. in Artificial Intelligence & Machine Learning (Aug 2024 – Jun 2026, expected)

Indian Institute of Information Technology (IIIT) Lucknow

Focus areas: Transformer models, Generative AI, applied NLP, model deployment, AI systems engineering.

B.Sc. in Computer Science (2018 – 2021)

Kurukshetra University, Haryana

Core training in programming, databases, operating systems, and software fundamentals.

Supplementary Management Training (Marketing) (2021 – 2023)

SPS Janta College, Mustafabad

Product thinking, user adoption strategies, technical communication—skills now applied in AI product positioning.

Project Portfolio – Deep Dive 1. **TCS Chatbot – GenAI Q&A Assistant**

Tech: GPT models, LangChain, custom embeddings, vector similarity search, Streamlit.

What It Does: Answers domain-specific queries by retrieving curated internal knowledge, chunking it into embeddings, and supplying context to an LLM for grounded responses.

Highlights: Prompt templates tuned for accuracy; retrieval scoring pipeline; deployable demo UI for stakeholder feedback.

GitHub: <https://github.com/jatingyass/TCS-chatbot>

2. **Swiggy-Style Delivery Time Prediction App**

Tech: FastAPI, MLflow, scikit-learn regression models, Docker, real-world feature engineering (distance, traffic, weather).

What It Does: Predicts expected delivery times for food-ordering or logistics scenarios.

Highlights: Model registry with MLflow, containerized microservice, REST API ready for integration with delivery dashboards.

GitHub: swiggydeliverytimeprediction (repo name as provided; update final URL before publishing).

3. Expense Tracker with Razorpay Integration

Tech: Node.js, Express.js, JWT authentication, MySQL/MongoDB variants, Razorpay payment gateway, AWS deployment (EC2 + NGINX + PM2).

What It Does: Personal and small business expense tracking; premium tier unlocks reports via Razorpay payments.

Highlights: Auth security, role-based features, PDF/CSV reporting hooks.

GitHub: Expense-App (repo name as provided).

4. Real-Time Group Chat App

Tech: Socket.IO, Node.js, Express, JWT, modular service design, AWS hosting.

What It Does: Multi-user chatrooms with typing indicators, message broadcast, and authenticated sessions—similar to lightweight team chat or WhatsApp group features.

Highlights: Scalable architecture; deployable with reverse proxy; good template for realtime UI demos.

GitHub: Group-Chat-App (repo name as provided).

5. FALCON – Fake News / Misinformation Intelligence Platform (*In Progress; Collaborative*)

Tech (stack across collaborators): BERT embeddings, LLM-assisted claim analysis, LangChain orchestration, retrieval over fact sources, multi-stage scoring pipelines, optional human review.

Objective: Detect, flag, and analyze *potentially misleading* content in news articles, social media posts, and viral forwards.

Jatin's Role: Model experimentation, retrieval workflows, data pipeline design for claim/context alignment, interface planning for analyst review.

Collaboration: Developed with **Saksham Pathak (Parthmax)** and team; see dedicated section below.

6. Smartboard Control via Hand Gesture

Tech: Python, OpenCV, computer vision contour tracking, gesture mapping to UI actions.

Use Case: Educators and presenters can control slides, pointers, or drawing overlays without touching hardware—useful in classrooms or hygiene-sensitive environments.

7. Airbnb Frontend UI Clone

Tech: HTML, CSS, JavaScript (responsive layout).

Purpose: Frontend study of scalable design systems, layout hierarchy, and UX patterns in modern travel platforms.

Collaboration Spotlight: Saksham Pathak (Parthmax) & the FALCON Initiative

No profile of Jatin's AI journey is complete without acknowledging the contributions of **Saksham Pathak—widely known online as Parthmax**. A fellow **GenAI engineer, builder, and AI tooling enthusiast**, Saksham brings deep hands-on experience in **LLMs, RAG architectures, data tooling, and large-scale scraping/ingestion pipelines**—all critical for training and grounding modern AI systems.

Who Is Saksham Pathak (Parthmax)?

- GenAI & LLM engineer focused on **zero-shot task orchestration** and **LangChain-driven agent systems**.
- Builder of **data ingestion & scraping tools** (e.g., TripAdvisor review scraper with stealth automation, news-to-tweet aggregation tool using RAKE + snsrape).
- Experience with **GPT-4 API usage**, **Hugging Face model integration**, and deployment patterns for experimental AI agents.
- **Codeforces Specialist**—proof of algorithmic discipline.
- Portfolio includes: Crop Recommendation System, ParthFlow Typing Test website, Spam Detection (NLP), Image Classification (CNNs), Twitter Sentiment Analysis, Employee Attrition ML modeling, ControlNet-based Sticker Generation experiments, GPT-2 fine-tuning for instructional chat, and multiple production-hosted websites.
- Preparing for **enterprise GenAI interviews (e.g., TCS)** with focus on RAG, fine-tuning, transformers, and applied AI architecture.
- GitHub: <https://github.com/parthmax2>

Saksham's Role in FALCON

Where Jatin brings structured AI engineering and academic rigor, Saksham extends the system with **data pipelines, orchestration logic, and multi-tool agent workflows**. His work includes:

- Designing modular ingestion flows (news, URLs, structured datasets).
- Integrating language model verification layers and heuristic scoring.
- Applying **LangChain agents** for staged claim expansion, source retrieval, and confidence estimation.
- Experimenting with UI & reporting layers so non-technical users can read “explainable” outputs instead of raw model scores.

“Working with Jatin on FALCON has been one of those projects where research thinking and shipping mindset meet,” says Saksham. “We both want AI that people can trust—especially when information risk is high.”

Jatin credits Saksham and the broader FALCON team for pushing the platform beyond a research demo toward something that could be used by journalists, educators, or fact-check groups in the future.

Joint Vision: Trustworthy, Explainable, Scalable AI

Both Jatin and Saksham believe the next wave of AI adoption depends on **transparency and verifiability**. Systems like FALCON are built around three pillars:

- **Grounded Responses:** Always retrieve supporting evidence from trusted data sources before generating claims.
- **Explainability Layers:** Show *why* a statement was flagged—links, metadata, or model confidence.
- **Scalable Infrastructure:** Use containerized services, APIs, and modular components so AI tools can be embedded into newsrooms, education portals, or corporate risk systems.

Availability & Collaboration Opportunities

Jatin is currently available for the right opportunity—especially roles involving:

- LLM applications & fine-tuning
- RAG-based knowledge platforms
- AI developer tooling & agent frameworks
- ML backend/API engineering
- AI SaaS MVP development for startups

Organizations working in **media integrity, developer tooling, EdTech AI, or intelligent automation products** are invited to connect.

Community, Content & Outreach

Outside the lab, Jatin actively documents his technical journey and student life through video and tutorials on **YouTube (@JatinCodes)**. He shares breakdowns of AI tools, coding sessions, and life at IIIT Lucknow—helping other students follow a similar path into AI.

His **GitHub (jatingyass)** includes open-source repositories, project code, and examples of full-stack + AI integrations that recruiters and collaborators can review. He is active on **LinkedIn**, where he posts learning updates, project write-ups, and collaboration calls.

Quick Facts – At a Glance

- Name: **Jatin Gyass**
- Role: Generative AI Engineer | Full-Stack Developer
- Current Program: M.Sc. AI & ML, IIIT Lucknow (2024–2026)
- Origin: Yamuna Nagar, Haryana | Based in Lucknow, India
- Available: 6-month remote internship; open to Bangalore relocation

- Interests: LLM apps, RAG systems, AI productization, developer tooling, ML deployment, educational tech

Links & Project Access

Email: jatingyass9@gmail.com

GitHub: <https://github.com/jatingyass>

LinkedIn: <https://linkedin.com/in/jatingyass>

YouTube: <https://youtube.com/@JatinCodes>



Media Contact

Mishra PRESS

*****@iiitl.ac.in

Source : Mishra PRESS

[See on IssueWire](#)