

# Muhammad Abdullah Ramzan

Faisalabad, PK · +92-306-9874952 · [abdullahramzan120@gmail.com](mailto:abdullahramzan120@gmail.com) · [linkedin.com/in/abdullahgrewal](https://www.linkedin.com/in/abdullahgrewal) · [buzzgrewal.com](https://www.buzzgrewal.com)

---

## EXPERIENCE

- 2025-Present **CHAWLA SOLUTIONS** Lahore, PK  
**Senior AI/ML Engineer - Hybrid**
- Designed and deployed end-to-end RAG-powered AI platforms and multi-agent systems, enabling intelligent automation and large-scale document interaction.
  - Developed high-performance RESTful APIs using Flask and FastAPI, supporting web and mobile clients with production-grade backend services.
  - Integrated ASR (Whisper, Deepgram) and TTS (Coqui, OpenAI) pipelines to create real-time, speech-to-speech conversational systems with natural interaction.
  - Engineered semantic search pipelines leveraging Sentence Transformers, ChromaDB, and PGVector, optimizing retrieval accuracy and efficiency for large document sets.
  - Led frontend and cloud deployment efforts using React, React Native, Tailwind CSS, and Azure with Docker + CI/CD, ensuring scalability, reliability, and seamless cross-platform performance.
- 2024-Present **MYNDEGE** Lahore, PK  
**Co-Founder & Lead AI Developer**
- **Built and scaled three investor-acquired AI products:**
    - AI-Powered Learning Platform Developed interactive, AI-driven courseware with integrated programming IDEs, real-time code evaluation, personalized learning paths, and generative AI content creation using React.js, Flask, LangChain, OpenAI APIs, and WebRTC for collaborative coding.
    - AI-Driven Maternity Care Chatbot Designed a medical consultancy assistant for expecting mothers, leveraging NLP (BERT, GPT-based LLMs), RAG pipelines, and medical knowledge graphs to provide safe, reliable guidance, integrated into mobile/web apps with Node.js, Firebase, and Flutter.
  - Led end-to-end product lifecycle from ideation and prototyping to deployment and investor handoff bridging business goals with AI-first engineering strategies.
- 2025-2025 **CODE DISTRICT** Lahore, PK  
**Associate AI/ML Engineer - Onsite**
- Spearheaded backend development using Python frameworks (Flask, FastAPI), delivering high-performance RESTful APIs integrated with web and mobile clients, improving system reliability by 45%.
  - Integrated AI features into existing Laravel projects, enhancing product capabilities with NLP, embeddings, and GPT-based automation, reducing manual workflows by 40%.
  - Collaborated with cross-functional teams to optimize AI model deployment pipelines, ensuring scalability, compliance, and low-latency inference across production environments.
- 2023-2025 **NEXOCORE** Islamabad, PK  
**Senior Full Stack AI Engineer (2024-2025)**
- Developed and deployed AI-powered web applications using Flask, React.js, Node.js, and SQL/NoSQL databases, integrating ML models (NLP, image processing, recommendation systems) via TensorFlow, PyTorch, and LangChain for production-ready APIs.
  - Deployed scalable, secure solutions on Azure and Heroku with Docker and CI/CD pipelines, optimizing performance through Redis caching, vector databases (FAISS, ChromaDB), and RAG-based architectures for intelligent search and automation.
- Associate Full Stack Engineer (2023-2024)**
- Assisted in developing and maintaining web applications using Flask, React.js, and SQL/NoSQL databases, contributing to feature implementation and bug fixes.
  - Collaborated with senior engineers to integrate basic AI/ML models and improve API performance, gaining hands-on experience in full-stack and AI-driven development.
- 

## SELECTED PROJECTS

- 2024-2025 **WISDOMWAVE** – [GITHUB](#)
- Built an AI-powered educational platform delivering personalized, textbook-based answers to enhance learning and collaboration for O/A Level students.
  - Integrated LLaMA 7B through Ollama for efficient LLM handling, with React.js frontend, Flask backend, ChromaDB for vector storage, and MongoDB for scalable data management. Implemented Retrieval-Augmented

Generation (RAG) with LangChain and llama.cpp to analyze textbooks and past papers, improving accuracy and contextual relevance of responses.

- Designed messaging and collaboration forums, enabling interactive student communities within the platform. Reduced query response latency by optimizing backend pipelines, improving student experience by 35%.

#### 2025-Present **CODE BLOOM – AI & EDU** – [GitHub](#)

- Built and deployed a scalable AI learning platform using Next.js, FastAPI, PostgreSQL (RDS), Prisma, and PQVector, featuring bilingual (English/Japanese) interaction and secure RBAC-based authentication.
- Implemented a RAG pipeline integrating OpenAI embeddings and GPT models for contextual Q&A over PDFs and YouTube transcripts, plus adaptive learning that personalizes lessons and quizzes based on student performance.
- Developed AI assistants (lesson chatbots, coding sandbox, project idea generator, and interview mentor) and an Admin Panel with analytics, caching (Redis), and Azure deployment achieving 99.9% uptime.

#### 2024-Present **MATERNAMIND – AI & MEDICAL** - [GitHub](#)

- Developed a multi-agent maternal healthcare platform using CrewAI with 4 specialized AI agents (Health, Care, Emotional, Analytics) powered by Mistral models and RAG orchestration, increasing user engagement by 65%.
- Implemented a medical RAG knowledge base with ChromaDB and OpenAI embeddings, delivering 92% accurate, evidence-based, and source-cited responses for pregnancy and mental health domains.
- Built end-to-end architecture featuring React + TypeScript (Tailwind) frontend and FastAPI + PostgreSQL backend with JWT/OAuth2, HIPAA compliance, and real-time analytics (vitals, fetal movement, PHQ-9/GAD-7/EPDS), reducing manual assessments by 40%.

#### 2024-2024 **MEDISYNAPSE – AI & MEDICAL** – [GitHub](#)

- Developed an AI-driven clinical assistant using Retrieval-Augmented Generation (RAG) on the MIMIC-IV-Ext dataset, achieving 91% contextual accuracy in clinical Q&A and patient summarization.
- Integrated multilingual voice-to-voice interaction with speech recognition, translation, and TTS, enabling natural, personalized medical conversations.
- Engineered a secure FastAPI backend with Hugging Face Transformers, HIPAA-compliant encryption, and optimized inference (vectorized retrieval + caching), reducing response latency by 38%.

---

## EDUCATION

2021-2025 **NATIONAL UNIVERSITY OF COMPUTER AND EMERGING SCIENCES - FAST NUCES** Islamabad, PK  
*Bachelor of Science in Computer Science; Major in Generative AI*

- **Awards:** Dean's List - Fall 2021, and Fall 2024 (top 10%)
- **Leadership:** Technical Head, FCAP (FAST Club of App Development & Programming) – leading 100+ members in major fields of CS, AI, GenAI, DS (2023 – 2025)
- **Teaching:** Teaching Assistant (Programming Fundamentals, OOP, Data Structures, Algorithms, Artificial Intelligence); supported 600+ students and 10+ faculty (2022 – 2025)

---

## HONORS & AWARDS

#### 2024-2025 **INTERNATIONAL COLLEGIATE PROGRAMMING CONTEST (ICPC)**

- Competed in the prestigious ICPC 2024–25, the world's largest and most renowned collegiate programming competition. Represented FAST NUCES at the regional level, tackling algorithmic, data structure, and optimization problems under strict time constraints.

#### 2025-Present **FAST PROBLEM-SOLVING COMPETITION (FPSC)**

- Achieved Top 3 ranking among 20+ competitive teams across the university. Successfully solved 5+ complex algorithmic problems, including graph theory, dynamic programming, and advanced data structures. Demonstrated strong analytical thinking, debugging under pressure, and speed coding abilities in a timed contest environment.

---

## LANGUAGES & TECHNICAL SKILL

- **Languages:** English (Fluent)
- **Coding Languages:** C++, C#, Python, SQL, JavaScript, PHP, SQL
- **Frameworks:** Flask, React.js, Express.js, Node.js, FAST API, Django, TensorFlow, Pytorch, LangChain, Laravel
- **Tools:** Git, XAMPP, Postman, Pandas, OpenCV, YOLO, Ollama, llama.cpp, Stripe API, ChromaDB, SQLite
- **Platforms:** Ubuntu, MongoDB, MySQL, RAG, XAMPP/WAMP