Personal Introduction

Name: Arooj Fatima
Role: AI/ML Engineer specializing in NLP, LLMs, and GenAI, with a passion for building scalable AI solutions like RAG systems and production-ready MLOps pipelines.

Professional Profiles

- **GitHub:** github.com/Aiengineer360 (visit Github for CV models, IR , Open source related projects)

- **Kaggle:** https://www.kaggle.com/ai360engineer

- **LinkedIn:**

- https://www.linkedin.com/in/arooj-fatima-04026b255?utm_source=share&utm_campaign=share_via&utm_content=profile&utm_medium=android_app

---

Key Sections to Include

1. **Education**

   o **Bachelor of Software Engineering (BSSE)**

      ▪ University of Management and Technology, Lahore (2021–2025) | CGPA: 3.6/4.0

   o **Pre-Medical (FSC)**

      ▪ Punjab Group of Colleges (PGC) | Score: 1053/1100 (96%)

   o **Matriculation**

      ▪ Siqarah Girls High School | Score: 1048/1100 (95%)

2. **Professional Experience**

   o **Data Scientist @ DataWorld (Jan 2025–Present)**

      ▪ Built time-series models (Facebook Prophet) for credit card spending predictions.

      ▪ Developed fraud detection systems using Isolation Forest and Z-scores.

      ▪ Analyzed FIFA and heart disease datasets to derive actionable insights.

   o **AI Engineer Intern @ GenZTechs (Oct–Dec 2024)**

      ▪ Implemented regression models for Boston housing price prediction.

      ▪ Engineered loan approval systems using feature selection techniques (Chi-Square, RFE).

3. **Technical Skills**

   o **Programming:** Python, SQL, React, TensorFlow

- AI/ML: NLP, LLMs (Gemini, OpenAI), Computer Vision, Time-Series Forecasting

- **Tools:** Hugging Face, Git, Docker, AWS, Streamlit

- **Certifications:**

    - Generative AI Engineer (Kaggle)

    - Google Data Scientist (Udemy)

    - IEEE Certifications (Innovative Lyallpur 3.0 FYP Exhibition)

4. **Projects**

    - **ShopSmart.AI**

        - Multi-agent shopping assistant using CrewAI, Serper API, and RAG.

        - Deployed via Streamlit on Render (Kaggle Competition).

    - **Legal Chatbot for Pakistani Laws**

        - Voice/text input (Whisper API), OpenAI, ChromaDB for legal judgments/drafting.

    - **Credit Card Fraud Detection**

        - Anomaly detection (Isolation Forest) and spending pattern analysis (Apriori, K-Means).

5. **Extracurricular & Interests**

    - **Martial Arts & Calligraphy**

        - Trained at Police Training College, Lahore: Self-defense, boxing, karate, shooting, and bike riding.

    - **Travel Enthusiast**

        - Explored Kashmir, Naran Kaghan, Murree, Swat, and more.

    - **Tech Community Engagement**

        - Attended **GenAI Xcelerate** (Systems Limited + Microsoft) to explore Copilot AI.

        - Participated in **Innovative Lyallpur 3.0** FYP exhibition (IEEE certified).

Academic Journey at UMT (2021–2025)

Bachelor of Software Engineering (BSSE)
CGPA: 3.60/4.00 | **Total Credit Hours Earned:** 117

Semester-wise Breakdown

1. **Fall 2021 (SGPA: 3.28)**

    - Programming Fundamentals (Lab: B)

- o English Grammar & Comprehension (A)
- o Islamic Studies (A)
- o Introduction to ICT (Lab: A)

2. **Spring 2022 (SGPA: 3.72)**
   - o Object-Oriented Programming (Lab: A)
   - o Discrete Structures (A)
   - o Applied Physics (Lab: A)
   - o Pakistan Studies (A)

3. **Fall 2022 (SGPA: 3.44)**
   - o Data Structures & Algorithms (Lab: B+)
   - o Software Engineering (A)
   - o Calculus & Linear Algebra (B)
   - o Psychology (A)

4. **Spring 2023 (SGPA: 3.34)**
   - o Database Systems (Lab: B+)
   - o Software Requirement Engineering (B+)
   - o Principles of Management (A)

5. **Fall 2023 (SGPA: 3.93)**
   - o Operating Systems (Lab: A)
   - o Computer Networks (A)
   - o Web Engineering (A)
   - o Foreign Language (A)

6. **Spring 2024 (SGPA: 3.72)**
   - o Information Security (A)
   - o Software Design & Architecture (Lab: A)
   - o Human-Computer Interaction (A)

7. **Fall 2024 (SGPA: 3.76)**
   - o **Natural Language Processing (A+)**
   - o **Information Retrieval (A+)**
   - o Software Construction (Lab: A+)
   - o Software Project Management (A+)

Key Technical Courses

- **AI/ML Focus:**
  - Natural Language Processing (A+)
  - Information Retrieval (A+)
  - Computer Vision

- **Software Engineering:**
  - Web Engineering (A)
  - Software Quality Engineering (B+)
  - Formal Methods in SE (A)

- **Core CS:**
  - Data Structures (A), Operating Systems (A), Databases (A)

Final Year Projects

- **CC4181/CC4182:** Final Year Project I & II (6 Cr. total)
  - *Upcoming:* Likely integrating NLP/GenAI (aligned with your professional skills).

Notable Achievements

- Consistent **A/A+ grades** in advanced courses (NLP, Web Eng, OS).
- **Highest SGPA (3.93)** in Fall 2023 with perfect scores in OS/Networking

Exploratory Data Analysis (EDA) , Data Science Projects

1. Credit Card Fraud Detection

Repo: GitHub Link
Dataset: Kaggle (284,807 transactions)

Key Insights:

- **Fraud Rate:** Only 0.17% of transactions were fraudulent (492 out of 284,807).

- **Transaction Analysis:**
  - Max legitimate transaction: **$25,691** (no fraud in high-value transactions).
  - Right-skewed amount distribution (most transactions < $5,000).

- **Technical Approach:**
  - Used **Isolation Forest** and **Z-scores** for anomaly detection.
  - Visualized skewness via histograms; avoided correlation heatmaps (low feature correlation).

2. Heart Disease Risk Analysis

Repo: [GitHub Link](#)
Dataset: 303 patients, 14 features (cholesterol, blood pressure, etc.)

Key Insights:

- **Risk Indicators:**

  - Avg. cholesterol **263 mg/dl** (with heart disease) vs. **223 mg/dl** (without).

  - High outliers in blood pressure/cholesterol (IQR detection).

- **Demographics:**

  - 75% patients male; age range 28–77.

  - Chest pain type 0 most common (496 cases).

- **Visualizations:**

  - Boxplots for outlier detection.

  - Histograms comparing health metrics by disease status.

---

3. FIFA Football Match Analysis

Repo: [GitHub Link](#)
Dataset: SQLite soccer data (matches, players, teams)

Key Insights:

- **Match Dynamics:**

  - Home teams win **more often** (goal difference analysis).

  - Teams with high **defensive pressure** and **aggression** win consistently.

- **Trends:**

  - Seasonal spikes in goals (specific months/tournaments).

- **Tools:**

  - Converted SQLite to CSV; used **Seaborn** for heatmaps.

Generative AI Projects Showcase

1. Kaggle Generative AI Course Projects

Repo: [GitHub](#) | **Certification:** [Kaggle](#)
Technologies: Python, Gemini API, LangGraph, Keras, TensorFlow

Key Modules:

- **RAG Pipeline:** Built a document Q&A system with retrieval-augmented generation.

- **Fine-Tuning:** Customized models for task-specific outputs.

- **Function Calling:** Integrated Gemini API for structured responses (e.g., JSON).

- **Evaluation:** Developed metrics for assessing model performance.

---

2. ShopSmart AI (Kaggle Competition)

Live Demo: ShopSmart.Ai | **Repo:** GitHub
Tech Stack: CrewAI, Gemini, RAG, Streamlit

Innovations:

- **Multi-Agent Workflow:**

  - *Search Agent:* Scrapes e-commerce platforms (e.g., Daraz).

  - *Review Agent:* Analyzes sentiment using text-embedding-004.

- **Voice Interface:** <500ms latency with audio preprocessing.

- **Deployment:** Auto-scaled on Render with 98% uptime.

Visual Highlight:

"User: 'Find budget laptops under $500'

→ Gemini validates query → CrewAI agents fetch & rank options → RAG summarizes top 3 picks."

---

3. MediBot - Medical Diagnosis Chatbot

Repo: GitHub
Tech Stack: GPT-4, Pinecone, Flask, AWS (EC2/ECR)

Key Features:

- **Symptom Analysis:** Differential diagnosis with confidence scoring.

- **Safety Checks:** Flags emergencies (e.g., chest pain → "Seek ER immediately").

- **Deployment:** CI/CD via GitHub Actions; Dockerized on AWS.

Blogging Website (Django)

Repo: GitHub | **Tech Stack:** Django, SQLite, HTML/CSS

Key Features:

- **User Auth:** Secure signup/login with email password reset.

- **Content Management:** Create/edit blog posts with categories/tags.

- **Interactive Elements:** Commenting system with user-level permissions.

- **Admin Panel:** Full CRUD control over users, posts, and comments.

Unique Aspect:

"Built a Medium-like platform where users retain full ownership of their content—implemented Django's class-based views for scalable post management."

AI Engineer Portfolio (Google Colab Style)

Live Demo: Portfolio | **Repo:** GitHub

Innovations:

- **UI Clone:** Mimics Google Colab's interface with interactive "code cells" displaying your skills/projects.

- **Dynamic Content:** Sections for certifications (clickable badges), AI projects with embedded demos.

- **Mobile-Friendly:** Responsive design with Material Icons.

Final Year Project: LegalAI-Pak – AI-Powered Legal Assistant

Live Demo: CourtCierge.Online

Project Vision

Mission: Democratize legal access in Pakistan through AI, bridging gaps for underserved communities.
Vision: Become Pakistan's leading AI-driven legal platform, making justice **understandable, affordable, and efficient**.

Key Features

1. **AI Legal Chatbot**

   o **NLP (GPT-4 Turbo + FAISS):** Answers queries in English about 25+ legal domains (e.g., Family, Criminal, Property Law).

   o **Document Drafting:** Generates notices, contracts, and petitions.

   o **Case Integration:** Pulls context from active cases for tailored advice.

2. **Case Workspace**

   o **Milestone Tracking:** Automated deadlines for hearings/filings with reminders.

   o **Secure Messaging:** E2E encrypted chats between clients/lawyers.

   o **Document Management:** Version-controlled storage (AWS S3) with role-based access.

3. **Lawyer Matchmaking**

   o **Verified Profiles:** BAR-licensed lawyers with success rates, specialties, and fee structures.

   o **Performance Analytics:** Win/loss ratios and case history (confidentiality-compliant).

4. **Community Q&A Forum**

   o Crowdsourced legal knowledge with responses vetted by professionals.

5. **Security & Compliance**

   o **GDPR/Pakistani Data Laws:** Encrypted data, MFA, and audit logs.

Tech Stack

Component	Technologies Used
Backend	Node.js, Express.js, JWT/OAuth
Database	MySQL (Sequelize ORM) + AWS RDS
Frontend	React, TypeScript, Tailwind, Socket.io
AI/ML	OpenAI GPT-4-turbo, FAISS vector embeddings

Target Users

- **General Public:** Free legal guidance.

- **Lawyers:** Case management tools.

- **NGOs/Corporates:** Compliance & contract support.

- **Students:** Research database.

Unique Innovations

- **Multilingual UI:** English/Urdu support (planned).

- **Automated Workflows:** Court date sync, doc generation.

- **Confidentiality:** Zero data retention policy.