Muhammad Ahmed

iahmedzahooor@gmail.com | +92 306 3398566 | LinkedIn | GitHub

SUMMARY

Software and AI Engineer with a strong focus on large language model evaluation, deep learning research, and building end to end mobile applications. I have partnered with top US firms like Turing and Revelo to benchmark frontier AI models, while also independently designing and publishing native mobile games to the Google Play Store. My background bridges machine learning optimization, robust backend databases, and consumer facing software.

EXPERIENCE

LLM Engineer (AI Evaluation) - Revelo (Feb 2026 to Present)

- Analyze how Anthropic's frontier models handle complex logic and follow strict instructions.
- Review AI-generated code changes in PR Writer workflows to ensure the output is safe, readable, and handles exceptions correctly.
- Run head to head model comparisons and provide targeted feedback that helps reduce model hallucinations.

LLM Projects - Turing (Nov 2024 to Jan 2026)

- Trained and benchmarked large language models across OpenAI, Anthropic, and Google families using RLHF.
- Tested output quality rigorously, focusing on multi step reasoning, truthfulness, and instruction following.
- Contributed to SEW Bench evaluations, video captioning tasks, and offline workflow improvements.
- Streamlined our data collection pipelines to process large datasets quickly while keeping our evaluation standards incredibly strict.
- Coordinated daily operations with international team members to consistently hit our model accuracy targets on time.

PROJECTS

2048 Game (Android) - Solo Developer [GitHub]

- Developed a complete 2048 game using Kotlin and Jetpack Compose, featuring a custom engine with 8 grid sizes and verified by 24 JUnit tests.

- Designed smooth gesture controls and stable animations, utilizing DataStore to save player settings and high scores locally.
- Linked Firebase Auth to real time Cloud Firestore leaderboards without relying on any external ad frameworks.
- Published a fully compliant, signed AAB build on Google Play.

PaintMaze Unity Mobile Puzzle Game - Solo Developer [GitHub] [Play Store]

- Created a mobile puzzle game in Unity and C# packed with 6,500 pre baked levels and on device generation for offline play.
- Engineered a custom board solver to generate hints and verify layouts, backed by a smart cache that eliminates loading screens.
- Integrated Firebase Auth for cloud saves and leaderboards, supported by automated Unity EditMode tests.

EDUCATION

BS Information Technology - Bahria University, Lahore (2023 to 2027)

SKILLS

- AI & Machine Learning: RLHF, Model Benchmarking, Multi Modal Annotation, Prompt Engineering, PyTorch, CNNs, Vision Transformers.
- Programming: Python, Kotlin, C#, JavaScript, C++, SQL.
- Mobile & Game Dev: Jetpack Compose, MVVM, StateFlow, DataStore, Android Studio, Unity.
- Backend & Infrastructure: PostgreSQL, SQL Server, Firebase Authentication, Cloud Firestore, Docker.
- Tools & Testing: Cursor, VS Code, Gradle, Kaggle, Graphviz, Apache JMeter.