+880 1321073452
adilshamim696@gmail.com

# ADIL SHAMIM

Machine Learning Engineer

Portfolio
adilshamim8
AdilShamim8

First-author researcher with a published paper achieving 56% inference speedup in Bengali speaker diarization (BUET CSE Fest 2026). **Kaggle** Top 1% globally (Rank 29/4,082). Production ML engineer who has shipped live systems serving real users—a recommendation engine delivering a verified +10% client sales lift, and **Toolly**, a live AI tools platform built and maintained solo. Specialized in Bengali NLP, production RAG systems, and end-to-end ML pipelines from training to deployment.

## PUBLICATIONS & RESEARCH

**Bangla Diarizz: Domain-Adapted Speaker Diarization for Bengali Long-Form Audio | 1st Author |** GitHub · HuggingFace Space **|** PyTorch · pyannote.audio · WeSpeaker

BUET CSE Fest 2026 — DL Sprint 4.0 Bengali Speaker Diarization Track | Rank #19 / 100+ teams

- Fine-tuned a segmentation model on a competition dataset; replaced speaker embeddings with WeSpeaker ResNet34-LM for targeted domain adaptation over language-agnostic baselines.
- Achieved 56% wall-clock inference speedup—reduced end-to-end inference from 1h 22m to ~36m across 14 test audio files.
- Achieved DER 0.286 on the private leaderboard. Deployed an interactive Gradio demo on HuggingFace Spaces with timeline visualization and downloadable RTTM output.

## EXPERIENCE

**Founder & Full-Stack ML Engineer — Toolly** Jun 2025 — Present

- Designed and shipped toolly.tech—a live AI tools directory with 400+ tools, 15 categories, a community submission pipeline, and an integrated Learn AI hub; all built and maintained solo.
- Engineered the full production stack: frontend, tool submission and moderation system, search and filter logic, and usage analytics from scratch.
- Built and deployed Toolly Studio — a Streamlit + Bria AI image generation app with batch export, Docker packaging, and a one-command demo flow for non-technical users.

**ML Engineer**—Independent Contractor | ALS Collaborative Filtering · Flask · Production Deployment | Jan -May 2025

- Built a hybrid recommendation engine (collaborative filtering ALS + content-based TF-IDF embeddings) for a local retail client; deployed as a Flask API to production.
- System delivered a verified +10% client sales lift within 90 days of deployment, quantified against pre/post client sales data.

## PROJECTS

**Production ML Pipeline — House Price Predictor |** GitHub | ZenML · MLflow · XGBoost · Docker · FastAPI

- Designed an end-to-end ML pipeline (ingest, preprocess, train, evaluate, register, serve) orchestrated with ZenML; integrated MLflow tracking for experiment versioning and hyperparameter logging.
- Implemented cross-validation and hyperparameter tuning in the training stage; deployed the final model as a Dockerized FastAPI inference service with structured logging and input validation.

**Production-grade-RAG |** GitHub | Production GenAI · LangChain · Qdrant · FastAPI · Inngest · OpenAI

- Production-ready RAG API with FastAPI, LangChain, OpenAI embeddings, and Qdrant vector store — covering chunking strategy, retrieval pipeline, and structured API response contracts.
- Implements query routing, context window management, and source attribution; designed for deployment behind a real inference endpoint with documented latency characteristics.

**Training Data Bot |** GitHub | LLM Engineering · Fine-Tuning Pipeline · PDF / URL Ingestion · Quality Scoring

- Automated pipeline that ingests raw documents (PDF, plain text, URLs), applies quality scoring, and outputs clean fine-tuning datasets formatted for LLM training — no manual curation step.
- Demonstrates full LLM engineering lifecycle thinking: data ingestion → preprocessing → quality filtering → structured output ready for fine-tuning runs.

## SKILLS

**Languages:** Python · SQL · Bash
**ML Frameworks:** PyTorch · TensorFlow · scikit-learn · XGBoost · LightGBM · CatBoost
**Deep Learning / NLP:** HuggingFace Transformers · Whisper · wav2vec2 · pyannote.audio · WeSpeaker · BERT
**MLOps & Deployment:** MLflow · ZenML · Docker · FastAPI · Flask · Streamlit · AWS · CI/CD · A/B testing
**Generative AI:** LLMs · RAG · LangChain · LangGraph · Qdrant · Prompt Engineering · Anthropic API · OpenAI API

## EDUCATION

B.Sc. Computer Science & Engineering — BNIST Feb 2023 — Present

- **Relevant coursework:** Linear Algebra · Calculus · Probability & Statistics · Data Structures & Algorithms · Operating Systems · Database Systems · System Design Projects · Artificial Intelligence · Machine Learning · Data Science

## Spoken Languages

- English (Professional — IELTS in progress) · Bengali (Native) · Hindi (Conversational)