MD. YEASIN ARAFAT

AI/ML & Full Stack Software Engineer

Computer Vision, OCR, LLMs & Data Quality | Full Stack Web Apps | Bangla-English Bilingual

Dhaka, Bangladesh | +880 1781173101 | codewithyeasin@gmail.com
linkedin.com/in/md-yeasin-arafat-b23484343 | github.com/CodeWithYeasin | codewithyeasin.github.io/yeasin-portfolio
Google Scholar: scholar.google.com/citations?user=xoYAoCcAAAAJ | ORCID: 0009-0004-4686-4362

Profile

Fourth-year Computer Science and Engineering student at North South University who builds AI/ML systems and the full-stack apps around them, with hands-on work in computer vision, semantic segmentation, OCR, LLMs and machine learning on real, imperfect datasets. Co-authored an IEEE conference paper (IS'26, published in IEEE Xplore) that benchmarks 9 LLMs on Bangla tasks with a blind human-rating rubric and cuts handwritten Bangla OCR error from 28% to 12% with QLoRA. Co-authored a research study on localizing AI-generated image manipulation, which meant curating a 9,611-image dataset with pixel-level masks and checking label quality down to the boundary. Comfortable across the full data lifecycle (preparing and labeling data, defining taxonomies, benchmarking against ground truth, structured error analysis) and on the product side with FastAPI, Django, React, Next.js and Docker. Native Bengali speaker with strong written English. Looking for AI/ML engineering, full stack software engineering or AI data training roles where careful evaluation decides whether a model ships.

Core Strengths

- **Annotation & Mask Preparation:** Built a 9,611-image forgery dataset with binary pixel-level masks, disjoint category splits to prevent leakage, and zero-mask handling for authentic images; also worked with character-level bounding boxes and a 7-class defect taxonomy.
- **Quality & Boundary Evaluation:** Judged labels and predictions using IoU, Dice, Boundary F1, Hausdorff95, Character Error Rate (CER), confusion matrices, and ROC/AUC, measuring edge precision and not only region overlap.
- **Cross-Source Consistency:** Evaluated across five datasets with differing manipulation sources, annotation styles, and image characteristics, and documented where label conventions break down.
- **Bangla Language Data:** Hand-wrote and line-annotated a 1,500+ page Bangla handwriting dataset; rated Bangla LLM outputs on a blind accuracy/script/naturalness rubric, and know the common conjunct, diacritic and script-mixing errors.
- **Hallucination & Grounding Checks:** Built a verifier that labels LLM answers as valid, weak, or hallucinated against retrieved source evidence.
- **Precision & Documentation:** Prompt design and output review with Google Gemini API; implemented ML algorithms from scratch and wrote reproducible evaluation pipelines that reward exact specification compliance.
- **Full Stack Delivery:** Ship models as usable apps: FastAPI and Django REST backends, React/Next.js/Vite frontends, Docker images, and live demos on Hugging Face Spaces (Reviva, Fabric Defect Detection).

Education

North South University **Jan 2022 - Present**
B.Sc. in Computer Science and Engineering (Expected Graduation: Apr 2027) *Dhaka, Bangladesh*
- **Coursework:** Machine Learning, Data Structures, Algorithms, Database Systems, Software Engineering, Operating Systems, Computer Networks, Web Technologies, Microprocessor Interfacing and Embedded Systems.

Bangladesh Navy School & College **Jun 2018 - Dec 2020**
Higher Secondary Certificate (Science), GPA: 5.00 / 5.00 *Dhaka, Bangladesh*

Model Academy **2018**
Secondary School Certificate (Science), GPA: 5.00 / 5.00 *Dhaka, Bangladesh*

Research Experience

Capstone Researcher, BhashaLLM (Senior Design Project, CSE 499A/B) **2025 - 2026**
Dept. of ECE, North South University; Supervisor: Dr. Mohammad Rashedur Rahman *Dhaka, Bangladesh*
Bangla LLM Adaptation & Handwritten OCR on Consumer Hardware | PyTorch, Transformers, PEFT/QLoRA, BitsAndBytes, Qwen-VL, Llama 3.2, FastAPI | github.com/CodeWithYeasin/BhashaLLM
- Published as a peer-reviewed paper at the 13th IEEE International Conference on Intelligent Systems (IS'26), now in IEEE Xplore (DOI: 10.1109/IS71416.2026.11727246); presented it in Varna, Bulgaria on 3 Sep 2026.
- Benchmarked 9 open LLMs (1.5B-12B; Qwen, Gemma, Mistral, Llama) on Bangla translation, summarization, OCR correction, and creative writing, all on a single 16 GB GPU via 4-bit NF4 quantization.
- Designed a blind human-evaluation rubric (accuracy, script correctness, naturalness) with an inter-rater agreement protocol; identified script confusion (Devanagari mixed into Bangla) in up to 23% of small-model outputs.
- Selected Llama 3.2-11B as the strongest model: over 99% correct Bangla script, 14/15 translation score, 0.52 ROUGE-1 on summarization, and 88% OCR-error correction with a 1% false-positive rate.
- Co-built a 1,500+ page handwritten Bangla dataset with line-level segmentation and transcriptions (released on Kaggle); QLoRA fine-tuning of a Qwen-VL OCR model cut Character Error Rate from 28% to 12% (88% character accuracy), in the same range as Google Cloud Vision's reported 13.89% on a comparable benchmark.
- QLoRA fine-tuned Qwen-2.5-1.5B for Bangla continued pre-training and answer grading (validation loss 0.018); served models offline through a CLI and web UI with on-demand adapter swapping.

- Demoed the system at booth 23 of NSU's Capstone Project Showcase, Season 19 (Apr 2026).

Undergraduate Researcher, Image Forensics **2026**
Dept. of ECE, North South University; Supervisor: Dr. Riasat Khan *Dhaka, Bangladesh*
Localization of Stable Diffusion Inpainting Using Edge Supervision, Feature Fusion and XAI | PyTorch, ConvNeXt-L, U-Net, GradCAM | github.com/CodeWithYeasin/Image-forgry-detection-and-localization

- Curated a 9,611-image subset of the Text-Guided Inpainting Forgery dataset with pixel-level binary masks, disjoint COCO categories across splits, and a controlled 1:2.5 real-to-fake ratio.
- Benchmarked 12 convolutional, transformer, and hybrid segmentation models; extended the strongest baseline with SRM/BayarConv noise and DCT frequency streams, SE fusion, edge-aware and deep supervision, and contrastive learning.
- Reached 99.44% pixel accuracy, 94.32% precision, 92.27% recall, 87.42% IoU, and 93.29% Dice, raising IoU over the ConvNeXt-L baseline from 0.8489 to 0.8742.
- Measured mask boundary quality with Boundary F1 and Hausdorff95, and cross-domain transfer on four external datasets (+64.3% relative IoU on inpainting-style forgeries).
- Applied GradCAM with faithfulness analysis; explanation-to-mask correlation rose from 0.41 to 0.79 on the source domain.

Publications

- S. A. Binaaf, **M. Y. Arafat**, A. S. Chaklader, and R. M. Rahman, "BhashaLLM: A QLoRA-Based Framework for Bangla Text Generation and Handwritten Character Recognition," in *Proc. 13th IEEE International Conference on Intelligent Systems (IS'26)*, Varna, Bulgaria, Sep. 3-5, 2026. Published in IEEE Xplore. DOI: 10.1109/IS71416.2026.11727246 (https://doi.org/10.1109/IS71416.2026.11727246).

Technical Skills

Languages: Python, JavaScript, TypeScript, C, C++, Java, PHP, SQL, HTML/CSS, Bash
AI/ML: Computer Vision, Semantic Segmentation, OCR, NLP, LLM Evaluation, QLoRA/LoRA Fine-Tuning, RAG & Hallucination Verification, Ensemble Learning, Gradient Boosting, Feature Engineering, Model Evaluation, Explainable AI (GradCAM), ML from Scratch, Image Restoration and Super-Resolution (GFPGAN, Real-ESRGAN), Speech-to-Text (Whisper)
Frameworks: PyTorch, TensorFlow, Hugging Face Transformers, PEFT, BitsAndBytes, FastAPI, Django REST Framework, Next.js, React, Vite, Tailwind CSS, Streamlit, Gradio
Data & Tools: NumPy, Pandas, OpenCV, Albumentations, FAISS, sentence-transformers, PostgreSQL, MySQL, MongoDB, Django ORM, Socket.io, Git/GitHub, GitHub Actions, pytest, Linux, Docker, Jupyter, Postman, Hugging Face Spaces
Media & AI Agents: FFmpeg, Demucs, CLAP, Remotion, Blender, edge-tts; Agent Skills for Claude Code, OpenAI Codex and Gemini CLI
Embedded: STM32 (Cortex-M3), PlatformIO, GPIO, ADC, I2C, Keypad Scanning, SWD Debugging
AI APIs: Google Gemini API, OpenAI API, Tesseract OCR, PyMuPDF

Projects

AI/ML & COMPUTER VISION

Fabric Defect Detection Using Ensemble ML Models (v2: SpecTexNet, in progress) **2025 - 2026**
Python, NumPy, PyTorch, ResNet50, ConvNeXt, Streamlit, Gradio | github.com/CodeWithYeasin/FabricDefectDetection | Demo: huggingface.co/spaces/codewithyeasin/fabric-defect-detection

- Implemented gradient boosting from scratch in NumPy (custom decision trees, pruning, regularization) without Scikit-Learn.
- Classified 7 defect categories at 93.25% test accuracy and 0.99 macro-average AUC; built an evaluation pipeline with confusion matrix, precision/recall/F1, and ROC analysis to isolate weak classes. Deployed a Streamlit UI with 23 ms inference.
- v2 (in progress): SpecTexNet, a CNN with a learnable multi-scale spectral-residual branch and spectral-guided attention, plus orientation-aware augmentation; audited v1 for data leakage and built a leakage-free protocol (perceptual-hash de-duplication, stratified group 5-fold CV, bootstrap confidence intervals, McNemar and Friedman tests) with CI unit tests.

Reviva - Open-Source AI Photo & Video Enhancer **2026**
Python, FastAPI, PyTorch, GFPGAN, Real-ESRGAN, FFmpeg, React, Vite, Tailwind CSS, Docker | github.com/CodeWithYeasin/Reviva | Demo: huggingface.co/spaces/codewithyeasin/reviva

- Built a web app that restores faces and upscales old or blurry photos 2x or 4x (GFPGAN + Real-ESRGAN) and enhances short videos frame by frame with FFmpeg, with a before/after comparison slider.
- Designed it to keep nothing: no accounts or database, files processed in a RAM-backed temp directory that is deleted once the response streams, `Cache-Control: no-store`, magic-byte upload validation, per-IP rate limits, and security headers.
- Packaged GPU and CPU Docker images and deployed a public demo on Hugging Face Spaces.

LLMS, RAG & DOCUMENT AI

DocuMind AI - Verified Multimodal Document Reasoning System **2026**
Python, FastAPI, FAISS, sentence-transformers, OpenAI API, PyMuPDF, Tesseract, Streamlit, Docker | github.com/CodeWithYeasin/DocuMind-AI-Verified-Multimodal-Document-Reasoning-System

- Built a retrieval-augmented QA pipeline (parse, retrieve, answer, verify) that ingests PDFs and scanned images into metadata-tagged chunks and retrieves evidence with FAISS semantic search.
- Wrote a custom verification module that scores how well each answer is grounded in retrieved evidence and labels it VALID, WEAK, or HALLUCINATED with a confidence score.
- Exposed upload and question endpoints through FastAPI with file-type whitelisting and size limits; added a Streamlit UI, unit tests for parser, retriever, and verifier, and Docker packaging.

AI Social Contract Advisor **2025**

Next.js, Django, PostgreSQL, Google Gemini API | github.com/CodeWithYeasin/AI_Social_Contract_Advisor

- Classified contract clauses and assigned low, medium, and high risk levels against a defined scoring rubric; extracted text from noisy PDFs and scans with Tesseract OCR and PyMuPDF.
- Added Bengali translation for local users; implemented authenticated REST APIs and file upload handling.

FULL STACK WEB APPLICATIONS

AI E-Commerce Store with Natural-Language Product Management **2026**

Next.js 15, TypeScript, MongoDB, Socket.io, Clerk, OpenAI GPT-4o | github.com/CodeWithYeasin/web-store

- Built a full-stack store with an admin assistant that turns plain-language instructions into structured product records, descriptions, and pricing suggestions via GPT-4o.
- Added real-time customer-admin chat with Socket.io, role-based admin access, social login, and automatic pre-order switching when stock runs out.

SweatFit AI - Personalized Fitness and Nutrition Chatbot **2024**

React, Django, PostgreSQL | github.com/CodeWithYeasin/SweatFit-AI

- Built an AI platform for workout, nutrition, and progress tracking with third-party calorie/meal APIs, JWT authentication, and a responsive analytics dashboard.

AI AGENTS & DEVELOPER TOOLS

Everything-videoEditing - AI Agent Skill for Video Editing **2026**

Python, FFmpeg, Whisper, Demucs, CLAP, OpenCV, Remotion, Blender, edge-tts | github.com/CodeWithYeasin/Everything-VideoEditing-Skill-EVS

- Wrote an Agent Skill that turns Claude Code, OpenAI Codex and Gemini CLI into a video editor: it analyzes the footage (speech, pauses, noise, faces, shaky shots, exposure), asks 2-4 multiple-choice questions, then edits.
- Drives every edit from a small JSON plan rendered by an FFmpeg engine: jump cuts from Whisper word timings, face-aware reframing, color grades, voice clean-up and music ducking, Remotion motion graphics, and exports for 9:16, 16:9 and 4:5 with an SRT file and thumbnail.
- Added word-by-word captions with correct Bangla text shaping, Bangla and English voice-over, an ad builder that renders 60/30/15/6 s versions, and one-command installers for Windows, Linux and macOS.

Loopify - Self-Verifying Task Loop Skill **2026**

Markdown, Shell, JavaScript | github.com/CodeWithYeasin/loopify

- Built a portable prompt skill that compiles a plain task into a goal, a verifier and a budget, runs work-and-check cycles until the verifier passes, and reports how strong that verification was (execution, rubric, or self-assessment).
- Kept the logic in one core file with thin adapters for Claude, ChatGPT, Gemini and Grok, plus a build script that packages it as a Claude skill.

EMBEDDED SYSTEMS

PLANT-X - Smart Plant Health Monitor (Embedded System) **2026**

C++, STM32F103C8T6, PlatformIO, I2C, ADC, DHT11, SSD1306 OLED, ST-LINK | CSE 331L, NSU (Team of 4) | github.com/CodeWithYeasin/PLANT-X

- Built an offline plant monitor on an STM32 Blue Pill reading soil moisture (12-bit ADC, calibrated to measured wet/dry endpoints), temperature and humidity (DHT11, rejecting corrupted frames), and ambient light (LDR divider), with rate-limited buzzer/LED alerts on threshold breaches; about 1,309 BDT in parts.
- Wrote about 1,250 lines of firmware: a keypad-driven state machine with scrolling menus, an all-sensor dashboard, a boot-time hardware self-check, and five OLED games (Snake, Tetris, Breakout, and more) on a 128x64 display; switched to ST-LINK SWD flashing after the USB bootloader failed on cloned boards.

Languages, Interests & Hobbies

Languages: Bengali (Native), English (Professional working proficiency; English-medium instruction at North South University)
Interests: Applied AI, Computer Vision, Data Quality and Annotation, OCR, NLP, Image Forensics, Full-Stack AI Applications, AI Agents and Developer Tools
Hobbies: Weight training at the gym, exploring new restaurants

References

Dr. Mohammad Rashedur Rahman *Supervisor, Senior Design Project (BhashaLLM)*
Professor, Dept. of ECE, North South University; Ph.D. in Computer Science, University of Calgary, Canada
Email: rashedur.rahman@northsouth.edu | Phone: +880 1731-767637 | Google Scholar

Dr. Riasat Khan *Supervisor, Image Forensics Research*
Associate Professor, Dept. of ECE, North South University; Ph.D. in ECE, New Mexico State University, USA
Email: riasat.khan@northsouth.edu | Phone: +880 2 55668200, Ext. 6382 | Google Scholar

Dr. Sumaiya Tabassum Nimi
Assistant Professor, Dept. of ECE, North South University; Ph.D. in Computer Science, University of Missouri-Kansas City, USA
Email: sumaiya.nimi@northsouth.edu | Phone: +880 2 55668200, Ext. 6383 | Google Scholar