

Khondker Salman Sayeed

Software Engineer @ IQVIA

Website  Github  LinkedIn  Scholar 

WORK EXPERIENCE

Software Engineer, IQVIA

JUNE 2024 - PRESENT

- Architecting and developing scalable web applications that integrate with our enterprise AI systems. Maintaining production ML pipelines using Azure AI services, AWS Bedrock and Sagemaker.
- Working on the development and deployment of IQVIA's in-house Large Language Models for generating expert domain-specific insights.
- Developing a Mixture-of-Agents based AI Assistant pipeline using LangGraph, CrewAI. Serving business users with data-informed intelligent and interactive responses.
- Increased a user-facing model's performance from 70% f1 to 90% f1 by moving from zero-shot to a RAG-based system.

Software Engineer, ResTechAI

JUNE 2024 - OCTOBER 2024 (Part-time)

- Leading the development of an easy-to-use, drag and drop ML platform that ingests complex data from diverse sources and provides intelligent insights.
- Designing a micro-service-based architecture using FastAPI, PostgreSQL, Docker, and Kubernetes.
- Integrating tools like Ollama, LiteLLM, AnythingLLM, and Windmill to enable users to create ML pipelines through a graphical UI.

Undergraduate Research Assistant, CSE BUET NLP Lab

APRIL 2023 - JUNE 2024

- Prepared a benchmark illusion dataset that tests the visual understanding capabilities of VLMs. (Published at **COLM'24**)
- Researching multimodal deep learning and Vision Language Models (VLM) for undergraduate thesis.
- Curated a large dataset of newspaper articles, images, and summaries to improve the abstractive summarization task by including image data along with the article for VLMs.

Machine Learning Intern, Northwest Power Generation Company

APRIL 2023 - MAY 2023

- Worked with the powerplant maintenance team to develop a Multi-Layer Perceptron and an LSTM model to analyze the powerplant performance level and failure prediction.
- Worked with the powerplant maintenance team to develop a Multi-Layer Perceptron and an LSTM model to analyze the powerplant performance level and failure prediction.

EDUCATIONAL BACKGROUND

Bachelor of Science in Computer Science and Engineering Bangladesh University of Engineering and Technology

APRIL 2019 - MAY 2024

- CGPA: 3.9/4.0
- University **Merit Scholarship** Recipient (4/6 Semesters)
- University **Dean's List** Scholarship Recipient (All Semesters)
- Member of IEEE Computer Society and Signal Processing Society

CONTACT

salkhon.github.io
salkhon050@gmail.com
khondkersalman.sayeed@iqvia.com

PROGRAMMING LANGUAGES

- Python
- C/C++
- Java
- Javascript/Typescript

TECHNOLOGIES

- ExpressJS, React, NextJS
- Flask, FastAPI
- TailwindCSS, MUI
- MySQL, MongoDB, PostgreSQL
- Redis
- Docker, Kubernetes
- AWS, GCP, Azure
- GitHub, GitLab, Jira, Confluence
- Pytorch, Tensorflow, Keras, CUDA

RESEARCH

IllusionVQA: A Challenging Optical Illusion Dataset for Vision Language Models

- Accepted at *COLM 2024*.
- [\[arXiv\]](#) [\[webpage\]](#) [\[open review\]](#)

RNA-DCGen: Dual Constrained RNA Sequence Generation with LLM-Attack

- Accepted at *MLSB@NeurIPS'24*.
- [\[bioRxiv\]](#)

Too Late to Train, Too Early To Use? A Study on Necessity and Viability of Low-Resource Bengali LLMs

- Under review at *COLING 2025*.
- [\[arXiv\]](#)










Ophthalmic Biomarker Detection Using Ensembled Vision Transformers -- Winning Solution to IEEE SPS VIP Cup 2023

- Champion solution presented at *ICIP 2023*.
- [\[arXiv\]](#)







REFERENCES

- AKM Ashikur Rahman
Professor at CSE, BUET
ashikur@cse.buet.ac.bd
Relation: Advisor
- Rifat Shahriyar
Professor at CSE, BUET
rifat@cse.buet.ac.bd
Relation: Thesis Supervisor

PROJECTS

-  **Ticketing** — A microservices webapp that lets users buy and sell tickets
- Consists of 6 independent services, coordinated by an event-based architecture.
 - Backend: NodeJS, MongoDB, NATS Streaming, Redis
 - Frontend: NextJS, TypeScript, TailwindCSS
 - Orchestration: Kubernetes
-  **Sub-C Compiler** — A compiler for a subset of the C programming language
- Tools: Flex (Lexer), GNU Bison (Parser)
 - Language: C, C++
-  **Mobile-Doc** — A telemedicine web application
- Backend: FastAPI, MongoDB, Redis, BigQuery
 - Frontend: React, MUI
 - Deployment: Google Cloud (Backend), Netlify (Frontend)
 - Link: [Pokedoc](#)
-  **Suntech** — An e-commerce web application
- Backend: Flask, OracleDB
 - Frontend: Jinja2, HTML, CSS, Javascript
-  **Ray-Tracing** — A rendering pipeline implementing ray tracing in OpenGL scenes
- Language: C++
 - Algorithms: Phong Lighting Model, Recursive Reflection
 - Tool: OpenGL
-  **Enhanced xv6 OS** — Copy-On-Write and Memory Page-Replacement in the xv6 OS
- Language: C
 - System Calls, Process Schedulers, Page Replacement
-  **twitt3r** — A Twitter clone using the t3 stack
- Framework: NextJS
 - Stack: T3 (Typescript, trpc, TailwindCSS)
 - Persistence: PlanetScale, Prisma ORM
 - Deployment: Vercel
 - Link: [twitt3r](#)
-  **DocumentQA** — A langchain LLM document chatbot
- Language: Python
 - Libraries: langchain, OpenAI
 - UI: Chainlit
 - Database: ChromaDB Vector Database
-  **IllusionVQA** — A Challenging Optical Illusion Dataset for Vision Language Models
- This work is a part of my undergraduate thesis
 - Published at **COLM'24**
 - Paper: [IllusionVQA: A Challenging Optical Illusion Dataset for Vision Language Models](#)
 - Website: [IllusionVQA](#)

AWARDS AND COMPETITIONS

- IEEE Upsilon Pi Epsilon Honor Society Award 2023** 
Winner
- Awarded by IEEE Computer Society
MARCH 2024
- Received award based on academic excellence and extracurricular activities in the field of computing
 - Awarded to only 4 applicants internationally
 - Received a \$1000 scholarship
- IEEE Video and Image Processing Cup 2023** 
Champion
- Hosted by IEEE Signal Processing Society, Georgia Institute of Technology
ICIP 2023, Kuala Lumpur, Malaysia
AUGUST - OCTOBER 2023
- Task: Biomarker Prediction from retina OCT scans
 - Model: MaxViT, EVA02-ViT, InceptionNet, ResNet
 - Competition Page: [2023 IEEE SPS Video and Image Processing \(VIP\) Cup](#)
 - Team Name: Synapse
 - Paper: [Ophthalmic Biomarker Detection Using Ensembled Vision Transformers -- Winning Solution to IEEE SPS VIP Cup 2023](#)
- DLsprint 2022** 
Champion
- Hosted by CSE BUET, Bengali.AI and Optimizely
JULY - SEPTEMBER 2022
- Task: Speech Transcription on the Bengali Common Voices Dataset
 - Model: wav2vec2
 - Kaggle: [DL Sprint - BUET CSE Fest 2022 | Kaggle](#)
 - Team Name: YellowKing, Account: [Salman Khondker](#)
 - Github: [DLsprint2022-Champion](#)
 - Paper: [Applying wav2vec2 for Speech Recognition on BN-CV](#)
- EEE DAY Datathon 2023** 
1st Runners Up
- Hosted by EEE BUET, Apurba Technologies Ltd.
FEBRUARY - MARCH 2023
- Task: Grammatical Error Detection on an in-house dataset
 - Model: T5 Transformer
 - Kaggle: [Apurba presents Bhashabhrom: EEE DAY 2023 Datathon | Kaggle](#)
 - Team Name: Schadenfreude, Account: [Salman Khondker](#)
 - GitHub Link: [Bangla-GED](#)
 - Paper: [Bangla Grammatical Error Detection Using T5 Transformer Model](#)
- Robi Datathon 3.0 2024** 
1st Runners Up
- Hosted by Robi Axiata Ltd, Huawei
APRIL - MAY 2024
- Task: Customer Purchase Behavior Prediction
 - Model: Random Forest Regressor, XGBoost
 - Kaggle Qualifier Round: [Problem 1](#), [Problem 2](#), [Problem 3](#)
 - Team Name: YellowKing
- AI For Bangla 2023** 
Honorable Mention
- Hosted by EBLICT, Bangladesh Computer Council
DECEMBER 2022 - MARCH 2023
- Task: First word-level Bangladeshi Sign Language Dataset
 - Model: 3DCNN, Image-RNN, STGCN
 - Host Page: [EBLICT Facebook](#)
 - Team Name: Silent Sentinels
 - GitHub Link: [BdSL40](#)