



MD SHOKOR A RAHAMAN

Senior AI Engineer

✉ md_17009796@utp.edu.my

☎ +60135758623



Jalan Teknologi 1, 32610, Pangsapuri
Seri Iskandar, Perak, Malaysia

Md Shokor A Rahaman is an AI researcher. With a master's degree, 10 journal publications, and nearly 250 citations, he brings over nine years of industry experience in Python, ML, NLP, and computer vision. He has deployed proprietary Frontier and open-source LLMs in cloud environments, integrated RAG and structured outputs, and built agentic workflows to optimize data analytics, automation, and system efficiency. Known for his innovative leadership and collaboration, Shokor bridges cutting-edge research with practical AI solutions.

Experience

Superstring Capital, Zhejiang, China - Senior AI Engineer

Sep 2023 - Present

Key Responsibilities and Achievements:

- Deployed proprietary Frontier and open-source LLMs in cloud environments via APIs/HuggingFace, leveraging RAG, structured outputs, and assistants.
- Applied a 5-step strategy (data curation, baseline modeling, Frontier fine-tuning) to exceed Frontier benchmarks, boosting efficiency and cost-effectiveness.
- Built scalable APIs, integrated LLMs on Modal, and constructed agentic workflows (planning, memory, multi-agent) to enhance performance, scalability, and NLP capabilities.

Truth , Sydney, Australia - Machine Learning Engineer

June 2023 - Aug 2023

Key Responsibilities and Achievements:

- Implemented transformer models like LLaMA, GPT-3 and GPT-4, alongside advanced image segmentation techniques, to bolster the accuracy of our document verification systems.
- Developed perspective transformation algorithms to enhance image processing capabilities.
- Leveraged TensorFlow, PyTorch, and Python for machine learning implementation and large-scale data processing.
- Utilized GPT-4 for generating responses and automating form filling, improving customer interaction and validation efficiency.
- Translated cutting-edge research into practical enhancements for our machine learning models.

WPDEVELOPER, Dhaka, Bangladesh – Data Scientist

Aug 2020 - May 2023

Applicant Tracking System (Resume Parsing & Resume Scoring) for Easy.Jobs:

- Designed a multi-layer neural network-based language model for Named Entity Recognition to identify categories such as Name, Email, Phone Number, Address, Degree, College Name, Results, Company, Designation, Duration, Soft Skills, and Technologies.
- Integrated the developed resume parser with existing ATS systems, CRMs, and software using NLP, Spacy, BERT, Regex, TensorFlow, Theano, Python, Flask, Named Entity Recognition (NER), and OCR.
- Deployed the Applicant Tracking System (ATS) to the Easy.jobs SaaS product, utilizing large-scale autoregressive language models (GPT3, GPT J, GPT Neo, BERT, XLNet, etc.) to enhance CV screening efficiency.
- Employed transfer-learning techniques with pre-trained word embeddings like Glove, fastText, BERT, ELMO, and Universal Sentence Encoder for text similarity tasks in resume scoring.

Intelligent Recommendation System for Templately.com:

- Developed an Intelligent Recommendation System to provide users with the most relevant templates.
- Deployed a recommendation engine for recommending templates on templately.com using Pandas, Matrix Factorization, Restricted Boltzmann Machines, Clustering and LSH, Nearest-neighbors, TF-IDF, Classifiers (e.g., ANN or Naive Bayes), and Tensor Factorization, which resulted in a 14% increase in average download size.

- Employed transfer-learning techniques with pre-trained word embeddings like Glove, fastText, BERT, ELMO, and Universal Sentence Encoder for text similarity tasks in resume scoring.

Intelligent Recommendation System for Templately.com:

- Developed an Intelligent Recommendation System to provide users with the most relevant templates.
- Deployed a recommendation engine for recommending templates on templately.com using Pandas, Matrix Factorization, Restricted Boltzmann Machines, Clustering and LSH, Nearest-neighbors, TF-IDF, Classifiers (e.g., ANN or Naive Bayes), and Tensor Factorization, which resulted in a 14% increase in average download size.

E-commerce Classifier Model for Shopify:

- Researched, prototyped, built features, and optimized (hyperparameter tuning) state-of-the-art machine learning and deep learning techniques like SVM, Logistic Regression, Random Forest regression, LSTM, CNN, etc., using scikit-learn, Keras, and TensorFlow on CPU/GPU for an E-commerce classifier model.
- Implemented CI/CD/CT pipelines for the E-commerce classifier model and automated the data science platform by building and managing data pipelines and extensions that connect NoSQL databases, APIs, Machine Learning Engines, and UIs.
- Collaborated with Google Cloud Platform (GCP) and Heroku to enhance the E-commerce classifier model.

Shopify String Matching Model:

- Utilized Data Structure Algorithm (DSA) with Pandas for efficient string matching on large datasets, reducing execution time from 15-18 seconds to 0.4 seconds.
- Extracted, transformed, and loaded data from various sources, and integrated data for machine learning projects.

Universiti Teknologi Petronas, Perak, Malaysia — Junior Data Scientist *July 2018 - Mar 2020*

Journal Papers

Corresponding/ 1st Author

1. "Optimization of the hydropower energy generation using Meta-Heuristic approaches: A review," Energy Rep., vol. 6, pp. 2230–2248, 2020. Published in Energy Reports Q2 (ISI Q2, Scopus Q1, IF = 3.56) DOI: 10.1016/j.egy.2020.08.009
2. "Artificial Intelligence approach to total organic carbon content prediction in shale gas reservoir using well logs: A review" (Published in International Journal of Innovative Computing, Information and Control (IJICIC)) (Q2 ISI: IF 0.44/ Q2 Scopus). DOI: 10.24507/ijicic.17.02.539
3. "A Modified Niching Crow Search Approach to Well Placement". Published in Energies MDPI (Q2 ISI IF = 2.7/ Q2 Scopus). DOI: 10.3390/en14040857
4. "Feature Selection-Based Artificial Intelligence Techniques for Estimating Total Organic Carbon from Well Logs,". Published in Journal of Physics: Conference Series Q3 (SCOPUS Q3, IF = 0.23) DOI: 10.1088/1742-6596/1529/4/042084
5. "Evaluation of Tree-Based Ensemble Learning Algorithms to Estimate Total Organic Carbon from Wireline Logs". (Accepted in International Journal of Innovative Computing, Information and Control (IJICIC)) (Q2 ISI: IF 0.44/ Q2 Scopus).

Co-author

6. "Modelling and optimization of microhardness of electroless Ni-P-TiO2 composite coating based on machine learning approaches and RSM". Published in Journal of Materials Research and Technology (ISI Q1: IF= 5.289) DOI: 10.1016/j.jmrt.2021.03.063
7. "Comparison of Machine Learning Classifiers for Accurate Prediction of Real-Time Stuck Pipe Incidents," Energies, vol. 13, no. 14, p. 3683, 2020. Published in Energies MDPI Q2 (ISI IF = 2.7) DOI: 10.3390/en13143683
8. Missing well log data handling in complex lithology prediction: An NIS apriori algorithm approach. Published in International Journal of Innovative Computing, Information and Control (IJICIC) (Q2 ISI: IF = 0.44). DOI: 10.24507/ijicic.16.03.1077

Conference Papers

1. Estcon ICER2020 Evaluation of TOC Using Hybridization of Deep Neural Networks and Genetic Algorithms.

Book Chapter

1. M. S. A. Rahaman and P. Vasant, "Artificial Intelligence Approach for Predicting TOC from Well Logs in Shale Reservoirs: A Review," in Deep Learning Techniques and Optimization Strategies in Big Data Analytics, IGI Global, 2020, pp. 46–77. DOI: 10.4018/978-1-7998-1192-3.ch004.

South Asian College Chittagong — Server Engineer & Lecturer

Jan 2015 - April 2018

Skills

Python | MySQL | MongoDB | Django | Flask | Seaborn | Matplotlib | Pandas | NumPy | AWS | NLP | Git | GCP | REST API | C | Dataiku | ETL | DB | Data Bricks | SQL | DAX | ML | DL | PySpark | TensorFlow | A/B test | Bitbucket | Circle CI | Kubeflow | MLflow | Grafana | Airflow | Data Modeling / Statistical Modeling | ETL | Machine Learning | Data Visualization | Image Segmentation | Perspective Transformation | OpenCV | Strong analytical skills | Comfortable dealing with numerical data | Strong team player | Transformer | LLaMA | GPT 3 & 4 | Generative AI

Education

Universiti Teknologi PETRONAS, Malaysia

Aug 2018 - March 2020

- Master of Science in Data Science (Research-Based)

International Islamic University Chittagong, Bangladesh

July 2010 - Aug 2014

- Bachelor of Engineering
- Major in Electrical and Electronics Engineering

Certificates

Problem Solving (Basic) - HackerRank

Credential Id: <https://www.hackerrank.com/certificates/9d4f7f78cea9>

Python (Basic) - HackerRank

Credential Id: <https://www.hackerrank.com/certificates/6bca5340577a>

SQL (Basic) - HackerRank

Credential Id: <https://www.hackerrank.com/certificates/462fcaa70cde>

SQL (Intermediate) - HackerRank

Credential Id: <https://www.hackerrank.com/certificates/938ada41a7b2>

SQL (Advanced) - HackerRank

Credential Id: <https://www.hackerrank.com/certificates/a5495beb0063>

Google Scholar Profile - 11 Publications with 180+ citations in ML/DL

<https://scholar.google.com/citations?user=Iqpq2h4AAAAJ&hl=en>