Md Omar Faruk Rokon

Senior Data Scientist

AdTech, Walmart Inc
+1 (951) 575 8556
mroko001@ucr.edu
in *ofrokon*

Professional Summary

Senior Data Scientist with extensive experience in deploying and managing advanced NLP models for retrieval and relevance in e-commerce. Proven track record in leading projects that leverage state-of-the-art techniques to enhance search relevancy and ad recommendation systems.

Research Interests

My research interests include Natural Language Processing (NLP), Information Retrieval, Machine Learning, Deep Learning, and their applications in E-commerce and Sponsored Search.

Education

2018–2022   **Ph.D. in Computer Science**, *University of California - Riverside*
Supervisor: Prof. Michalis Faloutsos

2012–2017   **B.Sc. in Computer Science and Engineering**, *Bangladesh University of Engineering and Technology (BUET)*
Supervisor: Prof. Md Monirul Islam

Experience

2022– Present   **Senior Data Scientist**, *Walmart Inc*, Sunnyvale, USA
- Led the development of Walmart's sponsored search retrieval model using a Two-Tower BERT model, including A/B testing and model deployment.
- Developed relevance models to guardrail the relevance of ad searches in Walmart, ensuring high-quality ad placements.
- Worked on GenAI/LLM models for various Walmart sponsored search use cases, leveraging state-of-the-art techniques to enhance search relevancy and ad recommendations.

June 2021 – Sept 2021   **Summer Research Intern** , *Los Alamos National Laboratory*
- Designed and evaluated a statistical learning algorithm using optimization techniques to reconstruct the dynamic state matrix of PMU.
- Proposed a systematic approach to reconstruct state matrix considering only observed nodes.
- Used JuMP package in Julia for constrained nonlinear program optimization.

2018–2022   **Research Assistant**, *University of California - Riverside*
- Created the largest malware source code database with 7504 malware using ML and NLP techniques.
- Designed an embedding approach to represent GitHub repos for similarity, classification, and clustering tasks.
- Proposed a graph-based approach to establish influence among repositories.

Publications

1. Md Omar Faruk Rokon, Weizhi Du, Zhaodong Wang, Musen Wen, "Next-Gen Sponsored Search: Crafting the Perfect Query with Inventory-Aware RAG (InvAwr-RAG)-Based GenAI". SIGIR Ecom 2024.

2. Md Omar Faruk Rokon, Andrei Simion, Weizhi Du, Musen Wen, Hong Yao, Kuang-chih Lee, "Enhancement of E-commerce Sponsored Search Relevancy with LLM". SIGIR Ecom 2024.

3   Jianghong Zhou, Weizhi Du, Md Omar Faruk Rokon, Zhaodong Wang, Jiaxuan Xu, Isha Shah, Kuang-chih Lee, Musen Wen, "Enhanced E-Commerce Attribute Extraction: Innovating with Decorative Relation Correction and LLAMA 2.0-Based Annotation". arXiv preprint arXiv:2312.06684 (2023).

4   Md Omar Faruk Rokon, Risul Islam, Md Rayhanul Masud, Michalis Faloutsos, "PIMan: A Comprehensive Approach for Establishing Plausible Influence among Software Repositories". International Conference on Advances in Social Network Analysis and Mining, 2022.

5   Md Omar Faruk Rokon, Pei Yan, Risul Islam, Michalis Faloutsos, "Repo2Vec: A Comprehensive Embedding Approach for Determining Repository Similarity". 37th International Conference on Software Maintenance and Evolution (ICSME), 2021.

6   Md Omar Faruk Rokon, Risul Islam, Ahmad Darki, Vagelis Papalexakis, Michalis Faloutsos, "SourceFinder: Finding Malware Source-Code from Publicly Available Repositories". 23rd International Symposium on Research in Attacks, Intrusions and Defenses (RAID), 2020.

7   Nishat Ara Tania, Md Rayhanul Masud, Md Omar Faruk Rokon, Qian Zhang, and Michalis Faloutsos. "Who is Creating Malware Repositories on GitHub and Why?" In Companion Proceedings of the ACM on Web Conference 2024, pp. 955-958. 2024.

8   Risul Islam, Ben Treves, Md Omar Faruk Rokon, Michalis Faloutsos, "LinkMan: Hyperlink-Driven Misbehavior Detection in Online Security Forums". International Conference on Advances in Social Network Analysis and Mining, 2021.

9   Risul Islam, Md Omar Faruk Rokon, Ahmad Darki, Michalis Faloutsos, "HackerScope: The Dynamics of a Massive Hacker Online Ecosystem". International Conference on Advances in Social Network Analysis and Mining, 2020.

10   Risul Islam, Md Omar Faruk Rokon, Vagelis Papalexakis, Michalis Faloutsos, "TenFor: A Tensor-Based Approach for Extracting Interesting "Events" from Security Forums". International Conference on Advances in Social Network Analysis and Mining, 2020.

11   Risul Islam, Md Omar Faruk Rokon, Vagelis Papalexakis, Michalis Faloutsos, "RecTen: A Recursive Low Rank Tensor Factorization Method to Discover Hierarchical Patterns in Multi-modal Data". International AAAI Conference on Web and Social Media, 2021.

Technical Skills

- **Programming Languages:** Python, C++, Java, R, SQL, JavaScript, MATLAB
- **Frameworks and Libraries:** TensorFlow, PyTorch, Scikit-learn, Hugging Face Transformers (BERT, GPT, LLaMA)
- **NLP and Text Analysis:** NLTK, spaCy, Gensim, TextBlob
- **Data Processing and Analysis:** Pandas, NumPy, PySpark, Hadoop
- **Cloud Services and Storage:** AWS, Google Cloud Platform (GCP), Google Cloud Storage (GCS)
- **Database Management:** MySQL, PostgreSQL, MongoDB, Cassandra
- **Big Data Technologies:** Apache Spark, Hadoop
- **Model Deployment and MLOps:** Docker, Kubernetes, MLflow
- **Version Control:** Git, GitHub, Bitbucket
- **Machine Learning Concepts:** Encoder-Decoder Models, Transformers

Awards and Recognition

2011   Champion, Math Olympiad, Dhaka, Bangladesh

2018   Dean's Fellowship Award, University of California, Riverside

Leadership Experience

2021–2022   Graduate Student Mentor, Blue Gold and Black Mentorship, University of California - Riverside

2020–2021   Member, International Student Concern Committee, University of California - Riverside