Raquib Bin Yousuf

Arlington, VA

✉ raquib@vt.edu   in linkedin.com/in/raquibyousuf

LIBRARIES

Pytorch
Keras
Tensorflow
Transformer
Huggingface
Spacy
scikit-learn
Pandas
GeoPandas
NetworkX
Shapely
NLTK
Numpy
Gensim

PROGRAMMING

Python
Java
C++
MATLAB
SQL
PHP
HTML
Docker
Bash
LaTeX

LANGUAGES

English
Bengali

Computer Science PhD student working in the intersection of natural language processing and machine learning, focused on exploring and enhancing the ability of large language models (LLMs) to perform complex reasoning in data-centric domains and practical applications.

EXPERIENCE

The Washington Post Engineering   **Data Science & ML Intern**
2023   *Washington, DC*

Worked with state-of-the-art Gen-AI systems and LLMs to improve personalized news experiences and business solutions for the Post, which resulted in a funded collaboration between the Virginia Tech and the Washington Post.

Virginia Tech   **Research Assistant**
2019-   *Arlington, VA*

Design, implementation, and deployment of applied machine learning and NLP systems, ranging from embeddings to transformer-based large language models, funded by agencies such as **DARPA, NSF**, and others, for:

- Improving language modeling with augmentation, regularization and fine-tuning for various applications
- Automated end-to-end information extraction system from large scale scholarly articles
- Information, Context and Expertise Preservation for Task-Oriented Recommendations
- Anomaly detection of data using NLP methods
- Stable isotope ratio analysis with Gaussian Process and data valuation/selection
- Forecasting Migration Patterns and Land Border Encounters from surrogate indicators

Eastern University   **Lecturer**
2018-   *Dhaka, Bangladesh*

Led advanced programming and digital logic design courses, managed labs, exams, student supervision, and key administrative responsibilities.

SELECTED PUBLICATIONS

- **Yousuf, R. B.**, Defelice, N., Sharma, M., Xu, S., & Ramakrishnan, N. (2024). LLM Augmentations to support Analytical Reasoning over Multiple Documents. **IEEE BigData 2024**
- **Yousuf, R. B.**, Biswas, S., Kaushal, K. K., Dunham, J., Gelles, R., Muthiah, S., ... & Ramakrishnan, N. (2022). Lessons from Deep Learning applied to Scholarly Information Extraction: What Works, What Doesn't, and Future Directions. Data-driven Science of Science at **ACM KDD 2022**
- Sharma, M., Muralidhar, N., Xu, S., **Yousuf, R. B.**, & Ramakrishnan, N. (2024). Information Guided Regularization for Fine-tuning Language Models. **COLM 2024**

EDUCATION

Virginia Tech   **Arlington, VA**
Doctor of Philosophy, Computer Science and Application   *2019-*

Virginia Tech   **Blacksburg, VA**
Master of Science, Computer Science and Application   *2019-2022*

Bangladesh University of Engineering and Technology (BUET)   **Dhaka, Bangladesh**
Bachelor of Science, Computer Science & Engineering   *2013-2017*